



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁷ : C12N 15/12, C07K 14/47, 16/18, G01N 33/566, C12Q 1/68, C12N 15/11, 15/62, A01K 67/027, A61K 38/00		A2	(11) International Publication Number: WO 00/58473
			(43) International Publication Date: 5 October 2000 (05.10.00)
(21) International Application Number: PCT/US00/08621		(72) Inventors; and	
(22) International Filing Date: 31 March 2000 (31.03.00)		(75) Inventors/Applicants (for US only): SHIMKETTS, Richard, A. [US/US]; 191 Leete Street, West Haven, CT 06516 (US); LEACH, Martin [GB/US]; 884 School Street, Webster, MA 01570 (US).	
(30) Priority Data:		(74) Agent: ELRIFI, Ivor, R.; Mintz, Levin, Cohn, Ferris, Glovsky and Popeo, P.C., One Financial Center, Boston, MA 02111 (US).	
60/127,607 31 March 1999 (31.03.99) US			
60/127,636 2 April 1999 (02.04.99) US			
60/127,728 5 April 1999 (05.04.99) US			
09/540,763 30 March 2000 (30.03.00) US			
(63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Applications		(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).	
US 60/127,607 (CIP)			
Filed on 31 March 1999 (31.03.99)			
US 60/127,636 (CIP)			
Filed on 2 April 1999 (02.04.99)			
US 60/127,728 (CIP)			
Filed on 5 April 1999 (05.04.99)			
US 09/540,763 (CIP)			
Filed on 30 March 2000 (30.03.00)			
(71) Applicant (for all designated States except US): CURAGEN CORPORATION [US/US]; 555 Long Wharf Drive, 11th Floor, New Haven, CT 06511 (US).		Published <i>Without international search report and to be republished upon receipt of that report.</i>	
(54) Title: NUCLEIC ACIDS INCLUDING OPEN READING FRAMES ENCODING POLYPEPTIDES; "ORFX"			
(57) Abstract			
<p>The present invention provides open reading frames ORFX, encoding isolated polypeptides, as well as polynucleotides encoding ORFX and antibodies that immunospecifically bind to ORFX or any derivative, variant, mutant, or fragment of the ORFX polypeptides, polynucleotides or antibodies. The invention additionally provides methods in which the ORFX polypeptide, polynucleotide and antibody are used in detection and treatment of a broad range of pathological states, as well as to other uses.</p>			

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav	TM	Turkmenistan
BF	Burkina Faso	GR	Greece		Republic of Macedonia	TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's	NZ	New Zealand		
CM	Cameroon		Republic of Korea	PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakhstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

NOVEL POLYNUCLEOTIDES AND POLYPEPTIDES ENCODED THEREBY

5

BACKGROUND OF THE INVENTION

The invention relates generally to nucleic acids and polypeptides encoded thereby, and methods of using these nucleic acids and polypeptides.

10

SUMMARY OF THE INVENTION

The invention is based in part on the discovery of nucleic acids that include open reading frames encoding novel polypeptides, and on the polypeptides encoded thereby. The nucleic acids and polypeptides are collectively referred to herein as "ORFX".

15

Accordingly, in one aspect, the invention provides an isolated nucleic acid molecule (SEQ ID NO:2*n*-1, wherein *n* is an integer between 1-3161), that encodes novel polypeptide, or a fragment, homolog, analog or derivative thereof. The nucleic acid can include, *e.g.*, a nucleic acid sequence encoding a polypeptide at least 85% identical to a polypeptide comprising the amino acid sequences of SEQ ID NO:2*n*, wherein *n* is an integer between 1-3161. The nucleic acid can be, *e.g.*, a genomic DNA fragment, or a cDNA molecule.

20

Also included in the invention is a vector containing one or more of the nucleic acids described herein, and a cell containing the vectors or nucleic acids described herein.

The invention is also directed to host cells transformed with a recombinant expression vector comprising any of the nucleic acid molecules described above.

25

In another aspect, the invention includes a pharmaceutical composition that includes an ORFX nucleic acid and a pharmaceutically acceptable carrier or diluent.

In a further aspect, the invention includes a substantially purified ORF polypeptide, *e.g.*, any of the ORFX polypeptides encoded by an ORFX nucleic acid, and fragments, homologs, analogs, and derivatives thereof. The invention also includes a pharmaceutical composition that includes a ORFX polypeptide and a pharmaceutically acceptable carrier or diluent.

5 In a still a further aspect, the invention provides an antibody that binds specifically to an ORFX polypeptide. The antibody can be, *e.g.*, a monoclonal or polyclonal antibody, and fragments, homologs, analogs, and derivatives thereof. The invention also includes a pharmaceutical composition including ORFX antibody and a pharmaceutically acceptable carrier or diluent. The invention is also directed to isolated antibodies that bind to an epitope on a
10 polypeptide encoded by any of the nucleic acid molecules described above.

The invention also includes kits comprising any of the pharmaceutical compositions described above.

The invention further provides a method for producing an ORFX polypeptide by providing a cell containing a ORFX nucleic acid, *e.g.*, a vector that includes a ORFX nucleic
15 acid, and culturing the cell under conditions sufficient to express the ORFX polypeptide encoded by the nucleic acid. The expressed ORFX polypeptide is then recovered from the cell. Preferably, the cell produces little or no endogenous ORFX polypeptide. The cell can be, *e.g.*, a prokaryotic cell or eukaryotic cell.

The invention is also directed to methods of identifying an ORFX polypeptide or nucleic
20 acids in a sample by contacting the sample with a compound that specifically binds to the polypeptide or nucleic acid, and detecting complex formation, if present.

The invention further provides methods of identifying a compound that modulates the activity of a ORFX polypeptide by contacting ORFX polypeptide with a compound and determining whether the ORFX polypeptide activity is modified.

25 The invention is also directed to compounds that modulate ORFX polypeptide activity identified by contacting a ORFX polypeptide with the compound and determining whether the compound modifies activity of the ORFX polypeptide, binds to the ORFX polypeptide, or binds to a nucleic acid molecule encoding a ORFX polypeptide.

In a another aspect, the invention provides a method of determining the presence of or
30 predisposition of an ORFX-associated disorder in a subject. The method includes providing a sample from the subject and measuring the amount of ORFX polypeptide in the subject sample.

The amount of ORFX polypeptide in the subject sample is then compared to the amount of ORFX polypeptide in a control sample. An alteration in the amount of ORFX polypeptide in the subject protein sample relative to the amount of ORFX polypeptide in the control protein sample indicates the subject has a tissue proliferation-associated condition. A control sample is preferably taken from a matched individual, *i.e.*, an individual of similar age, sex, or other general condition but who is not suspected of having a tissue proliferation-associated condition. Alternatively, the control sample may be taken from the subject at a time when the subject is not suspected of having a tissue proliferation-associated disorder. In some embodiments, the ORFX is detected using a ORFX antibody.

In a further aspect, the invention provides a method of determining the presence of or predisposition of an ORFX-associated disorder in a subject. The method includes providing a nucleic acid sample, *e.g.*, RNA or DNA, or both, from the subject and measuring the amount of the ORFX nucleic acid in the subject nucleic acid sample. The amount of ORFX nucleic acid sample in the subject nucleic acid is then compared to the amount of an ORFX nucleic acid in a control sample. An alteration in the amount of ORFX nucleic acid in the sample relative to the amount of ORFX in the control sample indicates the subject has a tissue proliferation-associated disorder.

In a still further aspect, the invention provides method of treating or preventing or delaying a ORFX-associated disorder. The method includes administering to a subject in which such treatment or prevention or delay is desired a ORFX nucleic acid, a ORFX polypeptide, or an ORFX antibody in an amount sufficient to treat, prevent, or delay a tissue proliferation-associated disorder in the subject.

Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Although methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, suitable methods and materials are described below. All publications, patent applications, patents, and other references mentioned herein are incorporated by reference in their entirety. In the case of conflict, the present specification, including definitions, will control. In addition, the materials, methods, and examples are illustrative only and not intended to be limiting.

Other features and advantages of the invention will be apparent from the following detailed description and claims.

DETAILED DESCRIPTION OF THE INVENTION

The invention provides novel polypeptides and nucleotides encoded thereby. The polynucleotides and their encoded polypeptides can be grouped according to the functions played by their gene products. Such functions include, structural proteins, proteins from which associated with metabolic pathways fatty acid metabolism, glycolysis, intermediary metabolism, calcium metabolism, proteases, and amino acid metabolism, etc.

Included in the invention are 3161 novel nucleic acid sequences and their encoded polypeptides. The sequences are collectively referred to as "ORFX nucleic acids" or ORFX polynucleotides" and the corresponding encoded polypeptide is referred to as a "ORFX polypeptide" or ORFX protein". For example, an ORFX nucleic acid according to the invention is a nucleic acid including an ORF1 nucleic acid, and an ORF polypeptide according to the invention is a polypeptide that includes the amino acid sequence of an ORF1 polypeptide. Unless indicated otherwise, "ORFX" is meant to refer to any of the ORF1-3161 sequences disclosed herein.

Table 1 provides a summary of the ORFX nucleic acids and their encoded polypeptides are summarized in Table 1. Nucleic acid sequences and polypeptide sequences for ORFX nucleic acids according to the invention is provided in the section of the specification entitled "Disclosed Sequences of ORFX Nucleic Acid and Polypeptide Sequences."

Column 1 of Table 1, entitled "ORF #", denotes an ORF number assigned to a nucleic acid containing an open reading frame according to the invention.

Column 2 of Table 1, entitled "Internal Identification number (Nucleic Acid Sequence Identification Number, Polypeptide Sequence Identification Number), provides an internal identification number for the indicated ORF, along with sequence identification numbers (SEQ ID NOs.) corresponding to the indicated ORF. In general, for an ORF n according to the invention (wherein n is any integer from 1 to 3161), a nucleic acid corresponding to the ORF is SEQ ID NO:2 n -1, and an amino acid sequence encoded by the ORF is SEQ ID NO:2 n . For example, a nucleic acid sequence corresponding to an ORF1 nucleic acid is SEQ ID NO:1, and a polypeptide sequence corresponding to an ORF1 polypeptide is SEQ ID NO:2. Similarly, a

nucleic acid sequence corresponding to an ORF4 nucleic acid is SEQ ID NO:7, and a polypeptide sequence corresponding to an ORF4 polypeptide is SEQ ID NO:8; a nucleic acid sequence corresponding to an ORF198 nucleic acid sequence is SEQ ID NO:395, and a polypeptide sequence corresponding to an ORF198 polypeptide is SEQ ID NO:396. Nucleic acid sequences and polypeptide sequences for ORFX nucleic acids according to the invention are provided in the section of the specification entitled "Disclosed Sequences of ORFX Nucleic Acid and Polypeptide Sequences."

Column 2 of Table 1, entitled "Protein Similarity", lists previously described proteins that are related to polypeptides encoded by the ORFs. Genbank identifiers for the previously described proteins are provided. These can be retrieved from <http://www.ncbi.nlm.nih.gov/>.

To determine similarity to previously described proteins, polypeptides encoded by ORFX DNA sequences were tested using the Framesearch Algorithm against a nonredundant version of the GenPept Database from NCBI/Genbank. DNA sequences that had a score of '90' or above (Framesearch algorithm score, Edelman et. al. GCG Genetics) to a known protein were selected. Open reading frames were extended beyond the region of the protein matched using standard DNA translation and codon tables. Novel proteins that lacked a protein match were translated against the standard genetic codons and proteins with an ORF at least 80 amino acids and containing a Methionine start are included in the Table.

Column 3 of Table 3, entitled "Protein Domains", lists previously described protein domains, designated by pfam entries, that are present in polypeptides encoded by the ORFs. Also included in column 3 are proteins in which these domains are present. The pfam entries can be retrieved from <http://pfam.wustl.edu/>. DNA sequences were translated in all six frames and tested using the Hmmer Algorithm against the Pfam Database (References to the algorithm and Pfam database can be found at <http://pfam.wustl.edu>). Translated DNA sequences that matched a protein domain entry in the Pfam database AND had a score of 7.5' were selected.

Column 4 of Table 3, entitled "Protein Classification", lists the type of classification assigned for the protein, based on its homology. Examples of proteins in the classification include the following proteins:

Amylases

Amylase is responsible for endohydrolysis of 1,4-alpha-glucosidic linkages in oligosaccharides and polysaccharides. Variations in amylase gene may be indicative of delayed maturation and of various amylase producing neoplasms and carcinomas.

5 **Amyloid**

The serum amyloid A (SAA) proteins comprise a family of vertebrate proteins that associate predominantly with high density lipoproteins (HDL). The synthesis of certain members of the family is greatly increased in inflammation. Prolonged elevation of plasma SAA levels, as in chronic inflammation, 15 results in a pathological condition, called amyloidosis, which affects the liver, kidney and spleen and which is characterized by the highly insoluble accumulation of SAA in these tissues. Amyloid selectively inhibits insulin-stimulated glucose utilization and glycogen deposition in muscle, while not affecting adipocyte glucose metabolism. Deposition of fibrillar amyloid proteins intraneuronally, as neurofibrillary tangles, extracellularly, as plaques and in blood vessels, is characteristic of both Alzheimer's disease and aged Down's syndrome. 15 Amyloid deposition is also associated with type II diabetes mellitus.

Angiopoeitin

Members of the angiopoietin/fibrinogen family have been shown to stimulate the generation of new blood vessels, inhibit the generation of new blood vessels, and perform several roles in blood clotting. This generation of new blood vessels, called angiogenesis, is also an 20 essential step in tumor growth in order for the tumor to get the blood supply it needs to expand. Variation in these genes may be predictive of any form of heart disease, numerous blood clotting disorders, stroke, hypertension and predisposition to tumor formation and metastasis. In particular, these variants may be predictive of the response to various antihypertensive drugs and chemotherapeutic and anti-tumor agents.

25 **Apoptosis-related proteins**

Active cell suicide (apoptosis) is induced by events such as growth factor withdrawal and toxins. It is controlled by regulators, which have either an inhibitory effect on programmed cell

death (anti-apoptotic) or block the protective effect of inhibitors (pro-apoptotic). Many viruses have found a way of countering defensive apoptosis by encoding their own anti-apoptosis genes preventing their target-cells from dying too soon. Variants of apoptosis related genes may be useful in formulation of anti-aging drugs.

5 **Cadherin, Cyclin, Polymerase, Oncogenes, Histones, Kinases**

Members of the cell division/cell cycle pathways such as cyclins, many transcription factors and kinases, DNA polymerases, histones, helicases and other oncogenes play a critical role in carcinogenesis where the uncontrolled proliferation of cells leads to tumor formation and eventually metastasis. Variation in these genes may be predictive of predisposition to any form
10 of cancer, from increased risk of tumor formation to increased rate of metastasis. In particular, these variants may be predictive of the response to various chemotherapeutic and anti-tumor agents.

Colony-stimulating factor-related proteins

Granulocyte/macrophage colony-stimulating factors are cytokines that act in
15 hematopoiesis by controlling the production, differentiation, and function of 2 related white cell populations of the blood, the granulocytes and the monocytes-macrophages.

Complement-related proteins

Complement proteins are immune associated cytotoxic agents, acting in a chain reaction to exterminate target cells to that were opsonized (primed) with antibodies, by forming a
20 membrane attack complex (MAC). The mechanism of killing is by opening pores in the target cell membrane. Variations in 20 complement genes or their inhibitors are associated with many autoimmune disorders. Modified serum levels of complement products cause edemas of various tissues, lupus (SLE), vasculitis, glomerulonephritis, renal failure, hemolytic anemia, thrombocytopenia, and arthritis. They interfere with mechanisms of ADCC (antibody dependent
25 cell cytotoxicity), severely impair immune competence and reduce phagocytic ability. Variants of complement genes may also be indicative of type I diabetes mellitus, meningitis neurological disorders such as nemaline myopathy, neonatal hypotonia, muscular disorders such as congenital myopathy and other diseases.

Cytochrome

The respiratory chain is a key biochemical pathway which is essential to all aerobic cells. There are five different cytochromes involved in the chain. These are heme bound proteins which serve as electron carriers. Modifications in these genes may be predictive of ataxia areflexia, dementia and myopathic and neuropathic changes in muscles. Also, association with various types of solid tumors.

Kinesins

Kinesins are tubulin molecular motors that function to transport organelles within cells and to move chromosomes along microtubules during cell division. Modifications of these genes may be indicative of neurological disorders such as Pick disease of the brain, tuberous sclerosis.

Cytokines, Interferon, Interleukin

Members of the cytokine families are known for their potent ability to stimulate cell growth and division even at low concentrations. Cytokines such as erythropoietin are cell-specific in their growth stimulation; erythropoietin is useful for the stimulation of the proliferation of erythroblasts. Variants in cytokines may be predictive for a wide variety of diseases, including cancer predisposition.

G-protein coupled receptors

G-protein coupled receptors (also called R7G) are an extensive group of hormones, neurotransmitters, odorants and light receptors which transduce extracellular signals by interaction with guanine nucleotide-binding (G) proteins. Alterations in genes coding for G-coupled proteins may be involved in and indicative of a vast number of physiological conditions. These include blood pressure regulation, renal dysfunctions, male infertility, dopamine associated cognitive, emotional, and endocrine functions, hypercalcemia, chondrodysplasia and osteoporosis, pseudohypoparathyroidism, growth retardation and dwarfism.

Thioesterases

Eukaryotic thiol proteases are a family of proteolytic enzymes which contain an active site cysteine. Catalysis proceeds through a thioester intermediate and is facilitated by a nearby histidine side chain; an asparagine completes the essential catalytic triad. Variants of thioester associated genes may be predictive of neuronal disorders and mental illnesses such as Ceroid Lipofficinos, Neuronal 1, Infantile, Santavuori disease and more.

The key to the molecule type is as follows:

	Abbrev:	Title:
10	amylase	amylase protein
	amylaseinhib	amylase inhibitor
	amyloid	amyloid protein
15	apoptosis	apoptosis associated protein
	apoptosisinhib	apoptosis inhibitors
	apoptosisrecep	apoptosis receptors
	ATPase_associated	ATPase associated protein
20	biotindp	biotin dependent enzyme/protein
	cadherin	cadherin protein
	calcium_channel	calcium channel protein
	carboxylase	carboxylase protein
	cathepsin	cathepsin/carboxypeptidases
	cathepsininhib	cathepsin/carboxypeptidase inhibitor
25	chloride_channel	chloride channel protein
	collagen	collagen
	complement	complement protein
	complementrecep	complement receptor protein
	complementinhib	complement inhibitor
30	csf	colony stimulating factor
	csfrecept	colony stimulating factor receptor
	cyclin	cyclin protein
	cyto450	cytochrome p450 protein
	cytochrome	cytochrome related protein
35	deaminase	deaminase
	dehydrogenase	dehydrogenase
	desaturase	desaturase
	dna_rna_bind	DNA/RNA binding protein/factor
	dna_rna_inhib	DNA/RNA binding protein/factor inhibitor
40	dynein	dynein

	elastase	elastase
	elastaseinhib	elastase inhibitor
	eph	EPH family of tyrosine kinases
	esterase	esterase
5	esteraseinhib	esterase inhibitor
	fgf	fibroblast growth factor
	fgfreceptor	fibroblast growth factor receptor
	gaba	GABA receptor
	glucoamylase	glucoamylase
10	glucoronidase	glucoronidase
	glycoprotein	glycoprotein
	Guanylyl	guanylylate cyclase
	helicase	helicase
	histone	histone
15	HOM	homologous
	homeobox	homeobox protein
	hydrolase	hydrolase
	hydroxysteroid	hydroxysteroid associated protein
	hypoxanthine	hypoxanthine associated protein
20	immunoglob	immunoglobulin
	immunoglobrecept	immunoglobulin receptor
	interferon	interferon
	interleukin	interleukin
	interleukinrecept	interleukin receptor
25	isomerase	isomerase
	isomeraseinhibitor	isomerase inhibitor
	isomerasereceptor	isomerase receptor
	kinase	kinase
	kinaseinhibitor	kinase inhibitor
30	kinasereceptor	kinase receptor
	kinesin	kinesin
	laminin	laminin associated protein
	lipase	lipase
	metallothionein	metallothionein
35	MHC	major histocompatibility complex
	misc_channel	miscellaneous channel
	ngf	nerve growth factor
	nuci_recpt	nuclear receptor
	nuclease	nuclease
40	oncogene	oncogene associated protein
	oxidase	oxidase
	oxygenase	oxygenase
	peptidase	peptidase
	peroxidase	peroxidase
45	phosphatase	phosphatase
	phosphataseinhib	phosphatase inhibitor

	phosphorylase	phosphorylase
	PIR	PIR DATABASE (release 56, 29-OCT-1998)
5	polymerase	polymerase
	potassium_channel	potassium channel protein
	prostaglandin	prostaglandin
	protease	protease
	proteaseinhib	protease inhibitor
10	reductase	reductase
	ribosomalprot	ribosomal associated protein
	RTR	EMBLDATABASE translated entries not to be incorporated into SWISS-PROT (20-JUL-1998)
15	SIM	similar
	SPTR	EMBL DATABASE translated entries to be incorporated into SWISS-PROT (20-JUL-1998)
20	struct	structural associated protein
	sulfotransferase	sulfotransferase
	SWP	SWISS-PROT DATABASE (release 18-OCT-1998)
	SWPN	SWISS-PROT Update (release 11-NOV-98)
25	synthase	synthase
	tgf	transforming growth factor
	tgfreceptor	transforming growth factor receptor
	thioesterase	thioesterase
	thiolase	thiolase
	tm7	seven transmembrane domain G-protein coupled receptor
30	tnf	necrosis factor receptor
	traffic	tumor necrosis factor
	tnfreceptor	tumor trafficking associated protein
	TRN	EMBL DATABASE translated entries update (20-JUL-1998)
35	transcriptfactor	transcription factor
	transferase	transferase
	transport	transport protein
	tubulin	tubulin
40	ubiquitin	ubiquitin
	unclassified	Protein not categorized into one of the aforementioned protein families
	water channel	water channel protein

Column 5 of Table 1, entitled, "Cells or Tissues in Which Gene is Expressed", denotes tissues, represented by five digit numbers, in which RNA homologous to the ORF nucleic acid sequences is present. Tissues or cells corresponding to the numbers are provided in Table 2.

ORFX nucleic acids, and their encoded polypeptides, according to the invention are useful in a variety of applications and contexts. For example, various ORFX nucleic acids and polypeptides according to the invention are useful, *inter alia*, as novel members of the protein families indicated in Table 1, and/or according to the presence of domains and sequence relatedness to previously described proteins as summarized in Table 1.

ORFX nucleic acids and polypeptides according to the invention can also be used to identify cell types listed in Table 1 for an indicated ORFX according to the invention. Additional utilities for ORFX nucleic acids and polypeptides according to the invention are disclosed herein.

ORFX Nucleic Acids

The novel nucleic acids of the invention include those that encode an ORFX or ORFX-like protein, or biologically active portions thereof. The nucleic acids include nucleic acids encoding polypeptides that include the amino acid sequence of one or more of SEQ ID NO:2*n*, wherein $n = 1$ to 3161. The encoded polypeptides can thus include, *e.g.*, the amino acid sequences of SEQ ID NO: 2, 4, 6, 8, 10, . . . , 6310, 6312, 6314, 6316, 6318, 6320, and/or 6322.

In some embodiments, a nucleic acid encoding a polypeptide having the amino acid sequence of one or more of SEQ ID NO:2*n* (wherein $n = 1$ to 3161) includes the nucleic acid sequence of any of SEQ ID NO:2*n*-1 (wherein $n = 1$ to 3161), or a fragment thereof. Additionally, the invention includes mutant or variant nucleic acids of any of SEQ ID NO:2*n*-1 (wherein $n = 1$ to 3161), or a fragment thereof, any of whose bases may be changed from the disclosed sequence while still encoding a protein that maintains its ORFX-like activities and physiological functions. The invention further includes the complement of the nucleic acid sequence of any of SEQ ID NO:2*n*-1 (wherein $n = 1$ to 3161), including fragments, derivatives,

analogs and homolog thereof. The invention additionally includes nucleic acids or nucleic acid fragments, or complements thereto, whose structures include chemical modifications.

Also included are nucleic acid fragments sufficient for use as hybridization probes to identify ORFX-encoding nucleic acids (*e.g.*, ORFX mRNA) and fragments for use as
5 polymerase chain reaction (PCR) primers for the amplification or mutation of ORFX nucleic acid molecules. As used herein, the term "nucleic acid molecule" is intended to include DNA molecules (*e.g.*, cDNA or genomic DNA), RNA molecules (*e.g.*, mRNA), analogs of the DNA or RNA generated using nucleotide analogs, and derivatives, fragments and homologs thereof. The nucleic acid molecule can be single-stranded or double-stranded, but preferably is
10 double-stranded DNA.

"Probes" refer to nucleic acid sequences of variable length, preferably between at least about 10 nucleotides (nt), 100 nt, or as many as about, *e.g.*, 6,000 nt, depending on use. Probes are used in the detection of identical, similar, or complementary nucleic acid sequences. Longer length probes are usually obtained from a natural or recombinant source, are highly specific and
15 much slower to hybridize than oligomers. Probes may be single- or double-stranded and designed to have specificity in PCR, membrane-based hybridization technologies, or ELISA-like technologies.

An "isolated" nucleic acid molecule is one that is separated from other nucleic acid molecules that are present in the natural source of the nucleic acid. Examples of isolated nucleic
20 acid molecules include, but are not limited to, recombinant DNA molecules contained in a vector, recombinant DNA molecules maintained in a heterologous host cell, partially or substantially purified nucleic acid molecules, and synthetic DNA or RNA molecules. Preferably, an "isolated" nucleic acid is free of sequences which naturally flank the nucleic acid (*i.e.*, sequences located at the 5' and 3' ends of the nucleic acid) in the genomic DNA of the organism
25 from which the nucleic acid is derived. For example, in various embodiments, the isolated ORFX nucleic acid molecule can contain less than about 50 kb, 25 kb, 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0.5 kb or 0.1 kb of nucleotide sequences which naturally flank the nucleic acid molecule in genomic DNA of the cell from which the nucleic acid is derived. Moreover, an "isolated" nucleic acid molecule, such as a cDNA molecule, can be substantially free of other cellular
30 material or culture medium when produced by recombinant techniques, or of chemical precursors or other chemicals when chemically synthesized.

A nucleic acid molecule of the present invention, *e.g.*, a nucleic acid molecule having the nucleotide sequence of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161), or a complement of any of this nucleotide sequence, can be isolated using standard molecular biology techniques and the sequence information provided herein. Using all or a portion of the nucleic acid sequence of any of SEQ ID NO:2*n*-1 (wherein *n*=1 to 3161) as a hybridization probe, ORFX nucleic acid sequences can be isolated using standard hybridization and cloning techniques (*e.g.*, as described in Sambrook *et al.*, eds., MOLECULAR CLONING: A LABORATORY MANUAL 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989; and Ausubel, *et al.*, eds., CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, New York, NY, 1993.)

A nucleic acid of the invention can be amplified using cDNA, mRNA or alternatively, genomic DNA, as a template and appropriate oligonucleotide primers according to standard PCR amplification techniques. The nucleic acid so amplified can be cloned into an appropriate vector and characterized by DNA sequence analysis. Furthermore, oligonucleotides corresponding to ORFX nucleotide sequences can be prepared by standard synthetic techniques, *e.g.*, using an automated DNA synthesizer.

As used herein, the term "oligonucleotide" refers to a series of linked nucleotide residues, which oligonucleotide has a sufficient number of nucleotide bases to be used in a PCR reaction. A short oligonucleotide sequence may be based on, or designed from, a genomic or cDNA sequence and is used to amplify, confirm, or reveal the presence of an identical, similar or complementary DNA or RNA in a particular cell or tissue. Oligonucleotides comprise portions of a nucleic acid sequence having about 10 nt, 50 nt, or 100 nt in length, preferably about 15 nt to 30 nt in length. In one embodiment, an oligonucleotide comprising a nucleic acid molecule less than 100 nt in length would further comprise at least 6 contiguous nucleotides of any of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161), or a complement thereof. Oligonucleotides may be chemically synthesized and may be used as probes.

In another embodiment, an isolated nucleic acid molecule of the invention comprises a nucleic acid molecule that is a complement of the nucleotide sequence shown in any of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161). In another embodiment, an isolated nucleic acid molecule of the invention comprises a nucleic acid molecule that is a complement of the nucleotide sequence shown in any of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161), or a portion of this nucleotide sequence. A nucleic acid molecule that is complementary to the nucleotide sequence shown in

is one that is sufficiently complementary to the nucleotide sequence shown in of any of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161) that it can hydrogen bond with little or no mismatches to the nucleotide sequence shown in of any of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161), thereby forming a stable duplex.

- 5 As used herein, the term “complementary” refers to Watson-Crick or Hoogsteen base pairing between nucleotides units of a nucleic acid molecule, and the term “binding” means the physical or chemical interaction between two polypeptides or compounds or associated polypeptides or compounds or combinations thereof. Binding includes ionic, non-ionic, Von der Waals, hydrophobic interactions, etc. A physical interaction can be either direct or indirect.
- 10 Indirect interactions may be through or due to the effects of another polypeptide or compound. Direct binding refers to interactions that do not take place through, or due to, the effect of another polypeptide or compound, but instead are without other substantial chemical intermediates.

- Moreover, the nucleic acid molecule of the invention can comprise only a portion of the
- 15 nucleic acid sequence of any of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161), e.g., a fragment that can be used as a probe or primer, or a fragment encoding a biologically active portion of ORFX. Fragments provided herein are defined as sequences of at least 6 (contiguous) nucleic acids or at least 4 (contiguous) amino acids, a length sufficient to allow for specific hybridization in the case of nucleic acids or for specific recognition of an epitope in the case of amino acids, respectively,
- 20 and are at most some portion less than a full length sequence. Fragments may be derived from any contiguous portion of a nucleic acid or amino acid sequence of choice. Derivatives are nucleic acid sequences or amino acid sequences formed from the native compounds either directly or by modification or partial substitution. Analogs are nucleic acid sequences or amino acid sequences that have a structure similar to, but not identical to, the native compound but
- 25 differs from it in respect to certain components or side chains. Analogs may be synthetic or from a different evolutionary origin and may have a similar or opposite metabolic activity compared to wild type.

- Derivatives and analogs may be full length or other than full length, if the derivative or analog contains a modified nucleic acid or amino acid, as described below. Derivatives or
- 30 analogs of the nucleic acids or proteins of the invention include, but are not limited to, molecules comprising regions that are substantially homologous to the nucleic acids or proteins of the

invention, in various embodiments, by at least about 70%, 80%, 85%, 90%, 95%, 98%, or even 99% identity (with a preferred identity of 80-99%) over a nucleic acid or amino acid sequence of identical size or when compared to an aligned sequence in which the alignment is done by a computer homology program known in the art, or whose encoding nucleic acid is capable of hybridizing to the complement of a sequence encoding the aforementioned proteins under stringent, moderately stringent, or low stringent conditions. See *e.g.* Ausubel, *et al.*, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, New York, NY, 1993, and below. An exemplary program is the Gap program (Wisconsin Sequence Analysis Package, Version 8 for UNIX, Genetics Computer Group, University Research Park, Madison, WI) using the default settings, which uses the algorithm of Smith and Waterman (Adv. Appl. Math., 1981, 2: 482-489, which is incorporated herein by reference in its entirety).

A "homologous nucleic acid sequence" or "homologous amino acid sequence," or variations thereof, refer to sequences characterized by a homology at the nucleotide level or amino acid level as discussed above. Homologous nucleotide sequences encode those sequences coding for isoforms of ORFX polypeptide. Isoforms can be expressed in different tissues of the same organism as a result of, for example, alternative splicing of RNA. Alternatively, isoforms can be encoded by different genes. In the present invention, homologous nucleotide sequences include nucleotide sequences encoding for a ORFX polypeptide of species other than humans, including, but not limited to, mammals, and thus can include, *e.g.*, mouse, rat, rabbit, dog, cat, cow, horse, and other organisms. Homologous nucleotide sequences also include, but are not limited to, naturally occurring allelic variations and mutations of the nucleotide sequences set forth herein. A homologous nucleotide sequence does not, however, include the nucleotide sequence encoding human ORFX protein. Homologous nucleic acid sequences include those nucleic acid sequences that encode conservative amino acid substitutions (see below) in any of SEQ ID NO:2*n* (wherein *n* = 1 to 3161) as well as a polypeptide having ORFX activity. Biological activities of the ORFX proteins are described below. A homologous amino acid sequence does not encode the amino acid sequence of a human ORFX polypeptide.

The nucleotide sequence determined from the cloning of the human ORFX gene allows for the generation of probes and primers designed for use in identifying the cell types disclosed and/or cloning ORFX homologues in other cell types, *e.g.*, from other tissues, as well as ORFX homologues from other mammals. The probe/primer typically comprises a substantially purified

oligonucleotide. The oligonucleotide typically comprises a region of nucleotide sequence that hybridizes under stringent conditions to at least about 12, 25, 50, 100, 150, 200, 250, 300, 350 or 400 or more consecutive sense strand nucleotide sequence of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161); or an anti-sense strand nucleotide sequence of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161);
5 or of a naturally occurring mutant of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161).

Probes based on the human ORFX nucleotide sequence can be used to detect transcripts or genomic sequences encoding the same or homologous proteins. In various embodiments, the probe further comprises a label group attached thereto, *e.g.*, the label group can be a radioisotope, a fluorescent compound, an enzyme, or an enzyme co-factor. Such probes can be
10 used as a part of a diagnostic test kit for identifying cells or tissue which misexpress a ORFX protein, such as by measuring a level of a ORFX-encoding nucleic acid in a sample of cells from a subject *e.g.*, detecting ORFX mRNA levels or determining whether a genomic ORFX gene has been mutated or deleted.

"A polypeptide having a biologically active portion of ORFX" refers to polypeptides
15 exhibiting activity similar, but not necessarily identical to, an activity of a polypeptide of the present invention, including mature forms, as measured in a particular biological assay, with or without dose dependency. A nucleic acid fragment encoding a "biologically active portion of ORFX" can be prepared by isolating a portion of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161), that encodes a polypeptide having a ORFX biological activity (biological activities of the ORFX
20 proteins are summarized in Table 1), expressing the encoded portion of ORFX protein (*e.g.*, by recombinant expression *in vitro*) and assessing the activity of the encoded portion of ORFX. For example, a nucleic acid fragment encoding a biologically active portion of ORFX can optionally include a domain as shown in Table 1, column 4.

25 ORFX variants

The invention further encompasses nucleic acid molecules that differ from the disclosed ORFX nucleotide sequences due to degeneracy of the genetic code. These nucleic acids thus encode the same ORFX protein as that encoded by the nucleotide sequence shown in SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161). In another embodiment, an isolated nucleic acid molecule of
30 the invention has a nucleotide sequence encoding a protein having an amino acid sequence shown in any of SEQ ID NO:2*n* (wherein *n* = 1 to 3161).

In addition to the human ORFX nucleotide sequence shown in any of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161), it will be appreciated by those skilled in the art that DNA sequence polymorphisms that lead to changes in the amino acid sequences of ORFX may exist within a population (*e.g.*, the human population). Such genetic polymorphism in the ORFX gene may exist among individuals within a population due to natural allelic variation. As used herein, the terms "gene" and "recombinant gene" refer to nucleic acid molecules comprising an open reading frame encoding a ORFX protein, preferably a mammalian ORFX protein. Such natural allelic variations can typically result in 1-5% variance in the nucleotide sequence of the ORFX gene. Any and all such nucleotide variations and resulting amino acid polymorphisms in ORFX that are the result of natural allelic variation and that do not alter the functional activity of ORFX are intended to be within the scope of the invention.

Moreover, nucleic acid molecules encoding ORFX proteins from other species, and thus that have a nucleotide sequence that differs from the human sequence of any of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161), are intended to be within the scope of the invention. Nucleic acid molecules corresponding to natural allelic variants and homologues of the ORFX cDNAs of the invention can be isolated based on their homology to the human ORFX nucleic acids disclosed herein using the human cDNAs, or a portion thereof, as a hybridization probe according to standard hybridization techniques under stringent hybridization conditions.

In another embodiment, an isolated nucleic acid molecule of the invention is at least 6 nucleotides in length and hybridizes under stringent conditions to the nucleic acid molecule comprising the nucleotide sequence of any of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161). In another embodiment, the nucleic acid is at least 10, 25, 50, 100, 250, 500 or 750 nucleotides in length. In another embodiment, an isolated nucleic acid molecule of the invention hybridizes to the coding region. As used herein, the term "hybridizes under stringent conditions" is intended to describe conditions for hybridization and washing under which nucleotide sequences at least 60% homologous to each other typically remain hybridized to each other.

Homologs (*i.e.*, nucleic acids encoding ORFX proteins derived from species other than human) or other related sequences (*e.g.*, paralogs) can be obtained by low, moderate or high stringency hybridization with all or a portion of the particular human sequence as a probe using methods well known in the art for nucleic acid hybridization and cloning.

As used herein, the phrase "stringent hybridization conditions" refers to conditions under which a probe, primer or oligonucleotide will hybridize to its target sequence, but to no other sequences. Stringent conditions are sequence-dependent and will be different in different circumstances. Longer sequences hybridize specifically at higher temperatures than shorter sequences. Generally, stringent conditions are selected to be about 5°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength, pH and nucleic acid concentration) at which 50% of the probes complementary to the target sequence hybridize to the target sequence at equilibrium. Since the target sequences are generally present at excess, at T_m , 50% of the probes are occupied at equilibrium. Typically, stringent conditions will be those in which the salt concentration is less than about 1.0 M sodium ion, typically about 0.01 to 1.0 M sodium ion (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30°C for short probes, primers or oligonucleotides (e.g., 10 nt to 50 nt) and at least about 60°C for longer probes, primers and oligonucleotides. Stringent conditions may also be achieved with the addition of destabilizing agents, such as formamide.

Stringent conditions are known to those skilled in the art and can be found in CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6. Preferably, the conditions are such that sequences at least about 65%, 70%, 75%, 85%, 90%, 95%, 98%, or 99% homologous to each other typically remain hybridized to each other. A non-limiting example of stringent hybridization conditions is hybridization in a high salt buffer comprising 6X SSC, 50 mM Tris-HCl (pH 7.5), 1 mM EDTA, 0.02% PVP, 0.02% Ficoll, 0.02% BSA, and 500 mg/ml denatured salmon sperm DNA at 65°C. This hybridization is followed by one or more washes in 0.2X SSC, 0.01% BSA at 50°C. An isolated nucleic acid molecule of the invention that hybridizes under stringent conditions to the sequence of any of SEQ ID NO:2*n*-1 (wherein $n = 1$ to 3161) corresponds to a naturally occurring nucleic acid molecule. As used herein, a "naturally-occurring" nucleic acid molecule refers to an RNA or DNA molecule having a nucleotide sequence that occurs in nature (e.g., encodes a natural protein).

In a second embodiment, a nucleic acid sequence that is hybridizable to the nucleic acid molecule comprising the nucleotide sequence of any of SEQ ID NO:2*n*-1 (wherein $n = 1$ to 3161), or fragments, analogs or derivatives thereof, under conditions of moderate stringency is provided. A non-limiting example of moderate stringency hybridization conditions are

hybridization in 6X SSC, 5X Denhardt's solution, 0.5% SDS and 100 mg/ml denatured salmon sperm DNA at 55°C, followed by one or more washes in 1X SSC, 0.1% SDS at 37°C. Other conditions of moderate stringency that may be used are well known in the art. See, *e.g.*, Ausubel *et al.* (eds.), 1993, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, NY, and
5 Kriegler, 1990, GENE TRANSFER AND EXPRESSION, A LABORATORY MANUAL, Stockton Press, NY.

In a third embodiment, a nucleic acid that is hybridizable to the nucleic acid molecule comprising the nucleotide sequence of any of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161), or fragments, analogs or derivatives thereof, under conditions of low stringency, is provided. A
10 non-limiting example of low stringency hybridization conditions are hybridization in 35% formamide, 5X SSC, 50 mM Tris-HCl (pH 7.5), 5 mM EDTA, 0.02% PVP, 0.02% Ficoll, 0.2% BSA, 100 mg/ml denatured salmon sperm DNA, 10% (wt/vol) dextran sulfate at 40°C, followed by one or more washes in 2X SSC, 25 mM Tris-HCl (pH 7.4), 5 mM EDTA, and 0.1% SDS at
15 50°C. Other conditions of low stringency that may be used are well known in the art (*e.g.*, as employed for cross-species hybridizations). See, *e.g.*, Ausubel *et al.* (eds.), 1993, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, NY, and Kriegler, 1990, GENE TRANSFER AND EXPRESSION, A LABORATORY MANUAL, Stockton Press, NY; Shilo and Weinberg, 1981, *Proc Natl Acad Sci USA* 78: 6789-6792.

Conservative mutations

In addition to naturally-occurring allelic variants of the ORFX sequence that may exist in the population, the skilled artisan will further appreciate that changes can be introduced by mutation into the nucleotide sequence of any of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161), thereby leading to changes in the amino acid sequence of the encoded ORFX protein, without altering the functional ability of the ORFX protein. For example, nucleotide substitutions
25 leading to amino acid substitutions at "non-essential" amino acid residues can be made in the sequence of any of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161). A "non-essential" amino acid residue is a residue that can be altered from the wild-type sequence of ORFX without altering the biological activity, whereas an "essential" amino acid residue is required for biological activity. For example, amino acid residues that are conserved among the ORFX proteins of the present
30 invention, are predicted to be particularly unamenable to alteration.

Amino acid residues that are conserved among members of an ORFX family members are predicted to be less amenable to alteration. For example, an ORFX protein according to the present invention can contain at least one domain (*e.g.*, as shown in Table 1) that is a typically conserved region in an ORFX family member. As such, these conserved domains are not likely to be amenable to mutation. Other amino acid residues, however, (*e.g.*, those that are not conserved or only semi-conserved among members of the ORFX family) may not be as essential for activity and thus are more likely to be amenable to alteration.

Another aspect of the invention pertains to nucleic acid molecules encoding ORFX proteins that contain changes in amino acid residues that are not essential for activity. Such ORFX proteins differ in amino acid sequence from any of any of SEQ ID NO:2*n* (wherein *n* = 1 to 3161), yet retain biological activity. In one embodiment, the isolated nucleic acid molecule comprises a nucleotide sequence encoding a protein, wherein the protein comprises an amino acid sequence at least about 75% homologous to the amino acid sequence of any of SEQ ID NO:2*n* (wherein *n* = 1 to 3161). Preferably, the protein encoded by the nucleic acid is at least about 80% homologous to any of SEQ ID NO:2*n* (wherein *n* = 1 to 3161), more preferably at least about 90%, 95%, 98%, and most preferably at least about 99% homologous to SEQ ID NO:2.

An isolated nucleic acid molecule encoding a ORFX protein homologous to the protein of any of SEQ ID NO:2*n* (wherein *n* = 1 to 3161) can be created by introducing one or more nucleotide substitutions, additions or deletions into the corresponding nucleotide sequence, *i.e.* SEQ ID NO:2*n*-1 for the corresponding *n*, such that one or more amino acid substitutions, additions or deletions are introduced into the encoded protein.

Mutations can be introduced into SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161) by standard techniques, such as site-directed mutagenesis and PCR-mediated mutagenesis. Preferably, conservative amino acid substitutions are made at one or more predicted non-essential amino acid residues. A "conservative amino acid substitution" is one in which the amino acid residue is replaced with an amino acid residue having a similar side chain. Families of amino acid residues having similar side chains have been defined in the art. These families include amino acids with basic side chains (*e.g.*, lysine, arginine, histidine), acidic side chains (*e.g.*, aspartic acid, glutamic acid), uncharged polar side chains (*e.g.*, glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), nonpolar side chains (*e.g.*, alanine, valine, leucine, isoleucine, proline,

phenylalanine, methionine, tryptophan), beta-branched side chains (*e.g.*, threonine, valine, isoleucine) and aromatic side chains (*e.g.*, tyrosine, phenylalanine, tryptophan, histidine). Thus, a predicted nonessential amino acid residue in ORFX is replaced with another amino acid residue from the same side chain family. Alternatively, in another embodiment, mutations can be introduced randomly along all or part of a ORFX coding sequence, such as by saturation mutagenesis, and the resultant mutants can be screened for ORFX biological activity to identify mutants that retain activity. Following mutagenesis of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161), the encoded protein can be expressed by any recombinant technology known in the art and the activity of the protein can be determined.

In one embodiment, a mutant ORFX protein can be assayed for (1) the ability to form protein:protein interactions with other ORFX proteins, other cell-surface proteins, or biologically active portions thereof, (2) complex formation between a mutant ORFX protein and a ORFX receptor, (3) the ability of a mutant ORFX protein to bind to an intracellular target protein or biologically active portion thereof; (*e.g.*, avidin proteins); (4) the ability to bind BRA protein; or (5) the ability to specifically bind an anti-ORFX protein antibody.

Antisense

Another aspect of the invention pertains to isolated antisense nucleic acid molecules that are hybridizable to or complementary to the nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161), or fragments, analogs or derivatives thereof. An "antisense" nucleic acid comprises a nucleotide sequence that is complementary to a "sense" nucleic acid encoding a protein, *e.g.*, complementary to the coding strand of a double-stranded cDNA molecule or complementary to an mRNA sequence. In specific aspects, antisense nucleic acid molecules are provided that comprise a sequence complementary to at least about 10, 25, 50, 100, 250 or 500 nucleotides or an entire ORFX coding strand, or to only a portion thereof. Nucleic acid molecules encoding fragments, homologs, derivatives and analogs of a ORFX protein of any of SEQ ID NO:2*n* (wherein *n* = 1 to 3161) or antisense nucleic acids complementary to a ORFX nucleic acid sequence of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161) are additionally provided.

In one embodiment, an antisense nucleic acid molecule is antisense to a "coding region" of the coding strand of a nucleotide sequence encoding ORFX. The term "coding region" refers to the region of the nucleotide sequence comprising codons which are translated into amino acid

residues (e.g., the protein coding region of a human ORFX that corresponds to any of SEQ ID NO:2n (wherein n = 1 to 3161)). In another embodiment, the antisense nucleic acid molecule is antisense to a "noncoding region" of the coding strand of a nucleotide sequence encoding ORFX. The term "noncoding region" refers to 5' and 3' sequences which flank the coding region that are not translated into amino acids (i.e., also referred to as 5' and 3' untranslated regions).

Given the coding strand sequences encoding ORFX disclosed herein (e.g., SEQ ID NO:2n-1 (wherein n = 1 to 3161)), antisense nucleic acids of the invention can be designed according to the rules of Watson and Crick or Hoogsteen base pairing. The antisense nucleic acid molecule can be complementary to the entire coding region of ORFX mRNA, but more preferably is an oligonucleotide that is antisense to only a portion of the coding or noncoding region of ORFX mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of ORFX mRNA. An antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides in length. An antisense nucleic acid of the invention can be constructed using chemical synthesis or enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (e.g., an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids, e.g., phosphorothioate derivatives and acridine substituted nucleotides can be used.

Examples of modified nucleotides that can be used to generate the antisense nucleic acid include: 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxymethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil,

3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine. Alternatively, the antisense nucleic acid can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (*i.e.*, RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

The antisense nucleic acid molecules of the invention are typically administered to a subject or generated *in situ* such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding a ORFX protein to thereby inhibit expression of the protein, *e.g.*, by inhibiting transcription and/or translation. The hybridization can be by conventional nucleotide complementarity to form a stable duplex, or, for example, in the case of an antisense nucleic acid molecule that binds to DNA duplexes, through specific interactions in the major groove of the double helix. An example of a route of administration of antisense nucleic acid molecules of the invention includes direct injection at a tissue site. Alternatively, antisense nucleic acid molecules can be modified to target selected cells and then administered systemically. For example, for systemic administration, antisense molecules can be modified such that they specifically bind to receptors or antigens expressed on a selected cell surface, *e.g.*, by linking the antisense nucleic acid molecules to peptides or antibodies that bind to cell surface receptors or antigens. The antisense nucleic acid molecules can also be delivered to cells using the vectors described herein. To achieve sufficient intracellular concentrations of antisense molecules, vector constructs in which the antisense nucleic acid molecule is placed under the control of a strong pol II or pol III promoter are preferred.

In yet another embodiment, the antisense nucleic acid molecule of the invention is an α -anomeric nucleic acid molecule. An α -anomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual β -units, the strands run parallel to each other (Gaultier *et al.* (1987) *Nucleic Acids Res* 15: 6625-6641). The antisense nucleic acid molecule can also comprise a 2'-o-methylribronucleotide (Inoue *et al.* (1987) *Nucleic Acids Res* 15: 6131-6148) or a chimeric RNA-DNA analogue (Inoue *et al.* (1987) *FEBS Lett* 215: 327-330).

Ribozymes and PNA moieties

Such modifications include, by way of nonlimiting example, modified bases, and nucleic acids whose sugar phosphate backbones are modified or derivatized. These modifications are

carried out at least in part to enhance the chemical stability of the modified nucleic acid, such that they may be used, for example, as antisense binding nucleic acids in therapeutic applications in a subject.

In still another embodiment, an antisense nucleic acid of the invention is a ribozyme.

5 Ribozymes are catalytic RNA molecules with ribonuclease activity that are capable of cleaving a single-stranded nucleic acid, such as an mRNA, to which they have a complementary region. Thus, ribozymes (*e.g.*, hammerhead ribozymes (described in Haselhoff and Gerlach (1988) *Nature* 334:585-591)) can be used to catalytically cleave ORFX mRNA transcripts to thereby inhibit translation of ORFX mRNA. A ribozyme having specificity for a ORFX-encoding
10 nucleic acid can be designed based upon the nucleotide sequence of a ORFX DNA disclosed herein (*i.e.*, SEQ ID NO:2*n*-1 (wherein $n = 1$ to 3161)). For example, a derivative of a Tetrahymena L-19 IVS RNA can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved in a ORFX-encoding mRNA. See, *e.g.*, Cech *et al.* U.S. Pat. No. 4,987,071; and Cech *et al.* U.S. Pat. No. 5,116,742.
15 Alternatively, ORFX mRNA can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules. See, *e.g.*, Bartel *et al.*, (1993) *Science* 261:1411-1418.

Alternatively, ORFX gene expression can be inhibited by targeting nucleotide sequences complementary to the regulatory region of the ORFX (*e.g.*, the ORFX promoter and/or
20 enhancers) to form triple helical structures that prevent transcription of the ORFX gene in target cells. See generally, Helene. (1991) *Anticancer Drug Des.* 6: 569-84; Helene. *et al.* (1992) *Ann. N.Y. Acad. Sci.* 660:27-36; and Maher (1992) *Bioassays* 14: 807-15.

In various embodiments, the nucleic acids of ORFX can be modified at the base moiety, sugar moiety or phosphate backbone to improve, *e.g.*, the stability, hybridization, or solubility of
25 the molecule. For example, the deoxyribose phosphate backbone of the nucleic acids can be modified to generate peptide nucleic acids (see Hyrup *et al.* (1996) *Bioorg Med Chem* 4: 5-23). As used herein, the terms "peptide nucleic acids" or "PNAs" refer to nucleic acid mimics, *e.g.*, DNA mimics, in which the deoxyribose phosphate backbone is replaced by a pseudopeptide backbone and only the four natural nucleobases are retained. The neutral backbone of PNAs has
30 been shown to allow for specific hybridization to DNA and RNA under conditions of low ionic strength. The synthesis of PNA oligomers can be performed using standard solid phase peptide

synthesis protocols as described in Hyrup *et al.* (1996) above; Perry-O'Keefe *et al.* (1996) *PNAS* 93: 14670-675.

PNAs of ORFX can be used in therapeutic and diagnostic applications. For example, PNAs can be used as antisense or antigene agents for sequence-specific modulation of gene expression by, *e.g.*, inducing transcription or translation arrest or inhibiting replication. PNAs of ORFX can also be used, *e.g.*, in the analysis of single base pair mutations in a gene by, *e.g.*, PNA directed PCR clamping; as artificial restriction enzymes when used in combination with other enzymes, *e.g.*, S1 nucleases (Hyrup B. (1996) above); or as probes or primers for DNA sequence and hybridization (Hyrup *et al.* (1996), above; Perry-O'Keefe (1996), above).

In another embodiment, PNAs of ORFX can be modified, *e.g.*, to enhance their stability or cellular uptake, by attaching lipophilic or other helper groups to PNA, by the formation of PNA-DNA chimeras, or by the use of liposomes or other techniques of drug delivery known in the art. For example, PNA-DNA chimeras of ORFX can be generated that may combine the advantageous properties of PNA and DNA. Such chimeras allow DNA recognition enzymes, *e.g.*, RNase H and DNA polymerases, to interact with the DNA portion while the PNA portion would provide high binding affinity and specificity. PNA-DNA chimeras can be linked using linkers of appropriate lengths selected in terms of base stacking, number of bonds between the nucleobases, and orientation (Hyrup (1996) above). The synthesis of PNA-DNA chimeras can be performed as described in Hyrup (1996) above and Finn *et al.* (1996) *Nucl Acids Res* 24: 3357-63. For example, a DNA chain can be synthesized on a solid support using standard phosphoramidite coupling chemistry, and modified nucleoside analogs, *e.g.*, 5'-(4-methoxytrityl)amino-5'-deoxy-thymidine phosphoramidite, can be used between the PNA and the 5' end of DNA (Mag *et al.* (1989) *Nucl Acid Res* 17: 5973-88). PNA monomers are then coupled in a stepwise manner to produce a chimeric molecule with a 5' PNA segment and a 3' DNA segment (Finn *et al.* (1996) above). Alternatively, chimeric molecules can be synthesized with a 5' DNA segment and a 3' PNA segment. See, Petersen *et al.* (1975) *Bioorg Med Chem Lett* 5: 1119-11124.

In other embodiments, the oligonucleotide may include other appended groups such as peptides (*e.g.*, for targeting host cell receptors *in vivo*), or agents facilitating transport across the cell membrane (see, *e.g.*, Letsinger *et al.*, 1989, *Proc. Natl. Acad. Sci. U.S.A.* 86:6553-6556; Lemaitre *et al.*, 1987, *Proc. Natl. Acad. Sci.* 84:648-652; PCT Publication No. W088/09810) or

the blood-brain barrier (see, *e.g.*, PCT Publication No. W089/10134). In addition, oligonucleotides can be modified with hybridization triggered cleavage agents (See, *e.g.*, Krol *et al.*, 1988, *BioTechniques* 6:958-976) or intercalating agents. (See, *e.g.*, Zon, 1988, *Pharm. Res.* 5: 539-549). To this end, the oligonucleotide may be conjugated to another molecule, *e.g.*, a peptide, a hybridization triggered cross-linking agent, a transport agent, a hybridization-triggered cleavage agent, etc.

ORFX polypeptides

The novel protein of the invention includes the ORFX-like protein whose sequence is provided in any of SEQ ID NO:2*n* (wherein *n* = 1 to 3161). The invention also includes a mutant or variant protein any of whose residues may be changed from the corresponding residue shown in FIG. 1 while still encoding a protein that maintains its ORFX-like activities and physiological functions, or a functional fragment thereof. For example, the invention includes the polypeptides encoded by the variant ORFX nucleic acids described above. In the mutant or variant protein, up to 20% or more of the residues may be so changed.

In general, an ORFX-like variant that preserves ORFX-like function includes any variant in which residues at a particular position in the sequence have been substituted by other amino acids, and further include the possibility of inserting an additional residue or residues between two residues of the parent protein as well as the possibility of deleting one or more residues from the parent sequence. Any amino acid substitution, insertion, or deletion is encompassed by the invention. In favorable circumstances, the substitution is a conservative substitution as defined above. Furthermore, without limiting the scope of the invention, positions of any of SEQ ID NO:2*n* (wherein *n* = 1 to 3161) may be substituted such that a mutant or variant protein may include one or more substitutions

The invention also includes isolated ORFX proteins, and biologically active portions thereof, or derivatives, fragments, analogs or homologs thereof. Also provided are polypeptide fragments suitable for use as immunogens to raise anti-ORFX antibodies. In one embodiment, native ORFX proteins can be isolated from cells or tissue sources by an appropriate purification scheme using standard protein purification techniques. In another embodiment, ORFX proteins are produced by recombinant DNA techniques. Alternative to recombinant expression, a ORFX

protein or polypeptide can be synthesized chemically using standard peptide synthesis techniques.

An "isolated" or "purified" protein or biologically active portion thereof is substantially free of cellular material or other contaminating proteins from the cell or tissue source from which the ORFX protein is derived, or substantially free from chemical precursors or other chemicals when chemically synthesized. The language "substantially free of cellular material" includes preparations of ORFX protein in which the protein is separated from cellular components of the cells from which it is isolated or recombinantly produced. In one embodiment, the language "substantially free of cellular material" includes preparations of ORFX protein having less than about 30% (by dry weight) of non-ORFX protein (also referred to herein as a "contaminating protein"), more preferably less than about 20% of non-ORFX protein, still more preferably less than about 10% of non-ORFX protein, and most preferably less than about 5% non-ORFX protein. When the ORFX protein or biologically active portion thereof is recombinantly produced, it is also preferably substantially free of culture medium, *i.e.*, culture medium represents less than about 20%, more preferably less than about 10%, and most preferably less than about 5% of the volume of the protein preparation.

The language "substantially free of chemical precursors or other chemicals" includes preparations of ORFX protein in which the protein is separated from chemical precursors or other chemicals that are involved in the synthesis of the protein. In one embodiment, the language "substantially free of chemical precursors or other chemicals" includes preparations of ORFX protein having less than about 30% (by dry weight) of chemical precursors or non-ORFX chemicals, more preferably less than about 20% chemical precursors or non-ORFX chemicals, still more preferably less than about 10% chemical precursors or non-ORFX chemicals, and most preferably less than about 5% chemical precursors or non-ORFX chemicals.

Biologically active portions of a ORFX protein include peptides comprising amino acid sequences sufficiently homologous to or derived from the amino acid sequence of the ORFX protein, *e.g.*, the amino acid sequence shown in SEQ ID NO:2 that include fewer amino acids than the full length ORFX proteins, and exhibit at least one activity of a ORFX protein. Typically, biologically active portions comprise a domain or motif with at least one activity of the ORFX protein. A biologically active portion of a ORFX protein can be a polypeptide which is, for example, 10, 25, 50, 100 or more amino acids in length.

A biologically active portion of a ORFX protein of the present invention may contain at least one of the above-identified domains conserved between the FGF family of proteins. Moreover, other biologically active portions, in which other regions of the protein are deleted, can be prepared by recombinant techniques and evaluated for one or more of the functional activities of a native ORFX protein.

In an embodiment, the ORFX protein has an amino acid sequence shown in any of SEQ ID NO:2*n* (wherein *n* = 1 to 3161). In other embodiments, the ORFX protein is substantially homologous to any of SEQ ID NO:2*n* (wherein *n* = 1 to 3161) and retains the functional activity of the protein of any of SEQ ID NO:2*n* (wherein *n* = 1 to 3161), yet differs in amino acid sequence due to natural allelic variation or mutagenesis, as described in detail below. Accordingly, in another embodiment, the ORFX protein is a protein that comprises an amino acid sequence at least about 45% homologous, and more preferably about 55, 65, 70, 75, 80, 85, 90, 95, 98 or even 99% homologous to the amino acid sequence of any of SEQ ID NO:2*n* (wherein *n* = 1 to 3161) and retains the functional activity of the ORFX proteins of the corresponding polypeptide having the sequence of SEQ ID NO:2*n* (wherein *n* = 1 to 3161).

Determining homology between two or more sequences

To determine the percent homology of two amino acid sequences or of two nucleic acids, the sequences are aligned for optimal comparison purposes (*e.g.*, gaps can be introduced in either of the sequences being compared for optimal alignment between the sequences). The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, then the molecules are homologous at that position (*i.e.*, as used herein amino acid or nucleic acid "homology" is equivalent to amino acid or nucleic acid "identity").

The nucleic acid sequence homology may be determined as the degree of identity between two sequences. The homology may be determined using computer programs known in the art, such as GAP software provided in the GCG program package. See, *Needleman and Wunsch* 1970 *J Mol Biol* 48: 443-453. Using GCG GAP software with the following settings for nucleic acid sequence comparison: GAP creation penalty of 5.0 and GAP extension penalty of 0.3, the coding region of the analogous nucleic acid sequences referred to above exhibits a

degree of identity preferably of at least 70%, 75%, 80%, 85%, 90%, 95%, 98%, or 99%, with the CDS (encoding) part of the DNA sequence shown in SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161).

The term "sequence identity" refers to the degree to which two polynucleotide or polypeptide sequences are identical on a residue-by-residue basis over a particular region of comparison. The term "percentage of sequence identity" is calculated by comparing two optimally aligned sequences over that region of comparison, determining the number of positions at which the identical nucleic acid base (*e.g.*, A, T, C, G, U, or I, in the case of nucleic acids) occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the region of comparison (*i.e.*, the window size), and multiplying the result by 100 to yield the percentage of sequence identity. The term "substantial identity" as used herein denotes a characteristic of a polynucleotide sequence, wherein the polynucleotide comprises a sequence that has at least 80 percent sequence identity, preferably at least 85 percent identity and often 90 to 95 percent sequence identity, more usually at least 99 percent sequence identity as compared to a reference sequence over a comparison region. The term "percentage of positive residues" is calculated by comparing two optimally aligned sequences over that region of comparison, determining the number of positions at which the identical and conservative amino acid substitutions, as defined above, occur in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the region of comparison (*i.e.*, the window size), and multiplying the result by 100 to yield the percentage of positive residues.

Chimeric and fusion proteins

The invention also provides ORFX chimeric or fusion proteins. As used herein, a ORFX "chimeric protein" or "fusion protein" includes a ORFX polypeptide operatively linked to a non-ORFX polypeptide. A "ORFX polypeptide" refers to a polypeptide having an amino acid sequence corresponding to ORFX, whereas a "non-ORFX polypeptide" refers to a polypeptide having an amino acid sequence corresponding to a protein that is not substantially homologous to the ORFX protein, *e.g.*, a protein that is different from the ORFX protein and that is derived from the same or a different organism. Within a ORFX fusion protein the ORFX polypeptide can correspond to all or a portion of a ORFX protein. In one embodiment, a ORFX fusion protein comprises at least one biologically active portion of a ORFX protein. In another embodiment, a ORFX fusion protein comprises at least two biologically active portions of a

ORFX protein. Within the fusion protein, the term "operatively linked" is intended to indicate that the ORFX polypeptide and the non-ORFX polypeptide are fused in-frame to each other. The non-ORFX polypeptide can be fused to the N-terminus or C-terminus of the ORFX polypeptide.

For example, in one embodiment a ORFX fusion protein comprises a ORFX polypeptide operably linked to the extracellular domain of a second protein. Such fusion proteins can be further utilized in screening assays for compounds that modulate ORFX activity (such assays are described in detail below).

In another embodiment, the fusion protein is a GST-ORFX fusion protein in which the ORFX sequences are fused to the C-terminus of the GST (i.e., glutathione S-transferase) sequences. Such fusion proteins can facilitate the purification of recombinant ORFX.

In yet another embodiment, the fusion protein is a ORFX protein containing a heterologous signal sequence at its N-terminus. For example, the native ORFX signal sequence can be removed and replaced with a signal sequence from another protein. In certain host cells (e.g., mammalian host cells), expression and/or secretion of ORFX can be increased through use of a heterologous signal sequence.

In another embodiment, the fusion protein is a ORFX-immunoglobulin fusion protein in which the ORFX sequences comprising one or more domains are fused to sequences derived from a member of the immunoglobulin protein family. The ORFX-immunoglobulin fusion proteins of the invention can be incorporated into pharmaceutical compositions and administered to a subject to inhibit an interaction between a ORFX ligand and a ORFX protein on the surface of a cell, to thereby suppress ORFX-mediated signal transduction *in vivo*. In one nonlimiting example, a contemplated ORFX ligand of the invention is an ORFX receptor. The ORFX-immunoglobulin fusion proteins can be used to modulate the bioavailability of a ORFX cognate ligand. Inhibition of the ORFX ligand/ORFX interaction may be useful therapeutically for both the treatment of proliferative and differentiative disorders, as well as modulating (e.g., promoting or inhibiting) cell survival. Moreover, the ORFX-immunoglobulin fusion proteins of the invention can be used as immunogens to produce anti-ORFX antibodies in a subject, to purify ORFX ligands, and in screening assays to identify molecules that inhibit the interaction of ORFX with a ORFX ligand.

A ORFX chimeric or fusion protein of the invention can be produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different polypeptide sequences are ligated together in-frame in accordance with conventional techniques, *e.g.*, by employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers that give rise to complementary overhangs between two consecutive gene fragments that can subsequently be annealed and reamplified to generate a chimeric gene sequence (see, for example, Ausubel et al. (eds.) CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, 1992). Moreover, many expression vectors are commercially available that already encode a fusion moiety (*e.g.*, a GST polypeptide). A ORFX-encoding nucleic acid can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the ORFX protein.

ORFX agonists and antagonists

The present invention also pertains to variants of the ORFX proteins that function as either ORFX agonists (mimetics) or as ORFX antagonists. Variants of the ORFX protein can be generated by mutagenesis, *e.g.*, discrete point mutation or truncation of the ORFX protein. An agonist of the ORFX protein can retain substantially the same, or a subset of, the biological activities of the naturally occurring form of the ORFX protein. An antagonist of the ORFX protein can inhibit one or more of the activities of the naturally occurring form of the ORFX protein by, for example, competitively binding to a downstream or upstream member of a cellular signaling cascade which includes the ORFX protein. Thus, specific biological effects can be elicited by treatment with a variant of limited function. In one embodiment, treatment of a subject with a variant having a subset of the biological activities of the naturally occurring form of the protein has fewer side effects in a subject relative to treatment with the naturally occurring form of the ORFX proteins.

Variants of the ORFX protein that function as either ORFX agonists (mimetics) or as ORFX antagonists can be identified by screening combinatorial libraries of mutants, *e.g.*, truncation mutants, of the ORFX protein for ORFX protein agonist or antagonist activity. In one

embodiment, a variegated library of ORFX variants is generated by combinatorial mutagenesis at the nucleic acid level and is encoded by a variegated gene library. A variegated library of ORFX variants can be produced by, for example, enzymatically ligating a mixture of synthetic oligonucleotides into gene sequences such that a degenerate set of potential ORFX sequences is expressible as individual polypeptides, or alternatively, as a set of larger fusion proteins (e.g., for phage display) containing the set of ORFX sequences therein. There are a variety of methods which can be used to produce libraries of potential ORFX variants from a degenerate oligonucleotide sequence. Chemical synthesis of a degenerate gene sequence can be performed in an automatic DNA synthesizer, and the synthetic gene then ligated into an appropriate expression vector. Use of a degenerate set of genes allows for the provision, in one mixture, of all of the sequences encoding the desired set of potential ORFX sequences. Methods for synthesizing degenerate oligonucleotides are known in the art (see, e.g., Narang (1983) *Tetrahedron* 39:3; Itakura *et al.* (1984) *Annu Rev Biochem* 53:323; Itakura *et al.* (1984) *Science* 198:1056; Ike *et al.* (1983) *Nucl Acid Res* 11:477).

Polypeptide libraries

In addition, libraries of fragments of the ORFX protein coding sequence can be used to generate a variegated population of ORFX fragments for screening and subsequent selection of variants of a ORFX protein. In one embodiment, a library of coding sequence fragments can be generated by treating a double stranded PCR fragment of a ORFX coding sequence with a nuclease under conditions wherein nicking occurs only about once per molecule, denaturing the double stranded DNA, renaturing the DNA to form double stranded DNA that can include sense/antisense pairs from different nicked products, removing single stranded portions from reformed duplexes by treatment with S1 nuclease, and ligating the resulting fragment library into an expression vector. By this method, an expression library can be derived which encodes N-terminal and internal fragments of various sizes of the ORFX protein.

Several techniques are known in the art for screening gene products of combinatorial libraries made by point mutations or truncation, and for screening cDNA libraries for gene products having a selected property. Such techniques are adaptable for rapid screening of the gene libraries generated by the combinatorial mutagenesis of ORFX proteins. The most widely used techniques, which are amenable to high throughput analysis, for screening large gene libraries typically include cloning the gene library into replicable expression vectors,

transforming appropriate cells with the resulting library of vectors, and expressing the combinatorial genes under conditions in which detection of a desired activity facilitates isolation of the vector encoding the gene whose product was detected. Recursive ensemble mutagenesis (REM), a new technique that enhances the frequency of functional mutants in the libraries, can be used in combination with the screening assays to identify ORFX variants (Arkin and Yourvan (1992) PNAS 89:7811-7815; Delgrave *et al.* (1993) Protein Engineering 6:327-331).

Anti-ORFX Antibodies

The invention further encompasses antibodies and antibody fragments, such as F_{ab} or $(F_{ab})_2$, that bind immunospecifically to any of the proteins of the invention.

An isolated ORFX protein, or a portion or fragment thereof, can be used as an immunogen to generate antibodies that bind ORFX using standard techniques for polyclonal and monoclonal antibody preparation. Full-length ORFX protein can be used. Alternatively, the invention provides antigenic peptide fragments of ORFX for use as immunogens. The antigenic peptide of ORFX comprises at least 4 amino acid residues of the amino acid sequence shown in any of SEQ ID NO:2*n* (wherein $n = 1$ to 3161). The antigenic peptide encompasses an epitope of ORFX such that an antibody raised against the peptide forms a specific immune complex with ORFX. The antigenic peptide may comprise at least 6 aa residues, at least 8 aa residues, at least 10 aa residues, at least 15 aa residues, at least 20 aa residues, or at least 30 aa residues. In one embodiment of the invention, the antigenic peptide comprises a polypeptide comprising at least 6 contiguous amino acids of any of SEQ ID NO:2*n* (wherein $n = 1$ to 3161).

In an embodiment of the invention, epitopes encompassed by the antigenic peptide are regions of ORFX that are located on the surface of the protein, *e.g.*, hydrophilic regions. As a means for targeting antibody production, hydropathy plots showing regions of hydrophilicity and hydrophobicity may be generated by any method well known in the art, including, for example, the Kyte Doolittle or the Hopp Woods methods, either with or without Fourier transformation. See, *e.g.*, Hopp and Woods, 1981, Proc. Nat. Acad. Sci. USA 78: 3824-3828; Kyte and Doolittle 1982, J. Mol. Biol. 157: 105-142, each incorporated herein by reference in their entirety.

As disclosed herein, an ORFX protein sequence of any of SEQ ID NO:2*n* (wherein $n = 1$ to 3161), or derivatives, fragments, analogs or homologs thereof, may be utilized as immunogens in the generation of antibodies that immunospecifically-bind these protein components. The term "antibody" as used herein refers to immunoglobulin molecules and

immunologically active portions of immunoglobulin molecules, i.e., molecules that contain an antigen binding site that specifically binds (immunoreacts with) an antigen, such as ORFX. Such antibodies include, but are not limited to, polyclonal, monoclonal, chimeric, single chain, F_{ab} and $F_{(ab)2}$ fragments, and an F_{ab} expression library. In a specific embodiment, antibodies to human ORFX proteins are disclosed. Various procedures known within the art may be used for the production of polyclonal or monoclonal antibodies to a ORFX protein sequence of any of SEQ ID NO:2*n* (wherein $n = 1$ to 3161) or derivative, fragment, analog or homolog thereof. Some of these proteins are discussed below.

For the production of polyclonal antibodies, various suitable host animals (e.g., rabbit, goat, mouse or other mammal) may be immunized by injection with the native protein, or a synthetic variant thereof, or a derivative of the foregoing. An appropriate immunogenic preparation can contain, for example, recombinantly expressed ORFX protein or a chemically synthesized ORFX polypeptide. The preparation can further include an adjuvant. Various adjuvants used to increase the immunological response include, but are not limited to, Freund's (complete and incomplete), mineral gels (e.g., aluminum hydroxide), surface active substances (e.g., lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, dinitrophenol, etc.), human adjuvants such as *Bacille Calmette-Guerin* and *Corynebacterium parvum*, or similar immunostimulatory agents. If desired, the antibody molecules directed against ORFX can be isolated from the mammal (e.g., from the blood) and further purified by well known techniques, such as protein A chromatography to obtain the IgG fraction.

The term "monoclonal antibody" or "monoclonal antibody composition", as used herein, refers to a population of antibody molecules that contain only one species of an antigen binding site capable of immunoreacting with a particular epitope of ORFX. A monoclonal antibody composition thus typically displays a single binding affinity for a particular ORFX protein with which it immunoreacts. For preparation of monoclonal antibodies directed towards a particular ORFX protein, or derivatives, fragments, analogs or homologs thereof, any technique that provides for the production of antibody molecules by continuous cell line culture may be utilized. Such techniques include, but are not limited to, the hybridoma technique (see Kohler & Milstein, 1975 *Nature* 256: 495-497); the trioma technique; the human B-cell hybridoma technique (see Kozbor, *et al.*, 1983 *Immunol Today* 4: 72) and the EBV hybridoma technique to produce human monoclonal antibodies (see Cole, *et al.*, 1985 In: MONOCLONAL ANTIBODIES AND

CANCER THERAPY, Alan R. Liss, Inc., pp. 77-96). Human monoclonal antibodies may be utilized in the practice of the present invention and may be produced by using human hybridomas (see Cote, *et al.*, 1983. *Proc Natl Acad Sci USA* 80: 2026-2030) or by transforming human B-cells with Epstein Barr Virus *in vitro* (see Cole, *et al.*, 1985 In: MONOCLONAL ANTIBODIES AND
5 CANCER THERAPY, Alan R. Liss, Inc., pp. 77-96). Each of the above citations are incorporated herein by reference in their entirety

According to the invention, techniques can be adapted for the production of single-chain antibodies specific to a ORFX protein (see *e.g.*, U.S. Patent No. 4,946,778). In addition, methods can be adapted for the construction of F_{ab} expression libraries (see *e.g.*, Huse, *et al.*,
10 1989 *Science* 246: 1275-1281) to allow rapid and effective identification of monoclonal F_{ab} fragments with the desired specificity for a ORFX protein or derivatives, fragments, analogs or homologs thereof. Non-human antibodies can be "humanized" by techniques well known in the art. See *e.g.*, U.S. Patent No. 5,225,539. Each of the above citations are incorporated herein by reference. Antibody fragments that contain the idiotypes to a ORFX protein may be produced by
15 techniques known in the art including, but not limited to: (i) an F_{(ab)₂} fragment produced by pepsin digestion of an antibody molecule; (ii) an F_{ab} fragment generated by reducing the disulfide bridges of an F_{(ab)₂} fragment; (iii) an F_{ab} fragment generated by the treatment of the antibody molecule with papain and a reducing agent and (iv) F_v fragments.

Additionally, recombinant anti-ORFX antibodies, such as chimeric and humanized
20 monoclonal antibodies, comprising both human and non-human portions, which can be made using standard recombinant DNA techniques, are within the scope of the invention. Such chimeric and humanized monoclonal antibodies can be produced by recombinant DNA techniques known in the art, for example using methods described in PCT International Application No. PCT/US86/02269; European Patent Application No. 184,187; European Patent
25 Application No. 171,496; European Patent Application No. 173,494; PCT International Publication No. WO 86/01533; U.S. Pat. No. 4,816,567; European Patent Application No. 125,023; Better *et al.* (1988) *Science* 240:1041-1043; Liu *et al.* (1987) *PNAS* 84:3439-3443; Liu *et al.* (1987) *J Immunol.* 139:3521-3526; Sun *et al.* (1987) *PNAS* 84:214-218; Nishimura *et al.* (1987) *Cancer Res* 47:999-1005; Wood *et al.* (1985) *Nature* 314:446-449; Shaw *et al.* (1988), *J.*
30 *Natl Cancer Inst* 80:1553-1559; Morrison (1985) *Science* 229:1202-1207; Oi *et al.* (1986) *BioTechniques* 4:214; U.S. Pat. No. 5,225,539; Jones *et al.* (1986) *Nature* 321:552-525;

Verhoeyan *et al.* (1988) *Science* 239:1534; and Beidler *et al.* (1988) *J Immunol* 141:4053-4060. Each of the above citations are incorporated herein by reference.

In one embodiment, methods for the screening of antibodies that possess the desired specificity include, but are not limited to, enzyme-linked immunosorbent assay (ELISA) and other immunologically-mediated techniques known within the art. In a specific embodiment, selection of antibodies that are specific to a particular domain of a ORFX protein is facilitated by generation of hybridomas that bind to the fragment of a ORFX protein possessing such a domain. Antibodies that are specific for one or more domains within a ORFX protein, *e.g.*, the domain spanning the first fifty amino-terminal residues specific to ORFX when compared to FGF-9, or derivatives, fragments, analogs or homologs thereof, are also provided herein.

Anti-ORFX antibodies may be used in methods known within the art relating to the localization and/or quantitation of a ORFX protein (*e.g.*, for use in measuring levels of the ORFX protein within appropriate physiological samples, for use in diagnostic methods, for use in imaging the protein, and the like). In a given embodiment, antibodies for ORFX proteins, or derivatives, fragments, analogs or homologs thereof, that contain the antibody derived binding domain, are utilized as pharmacologically-active compounds [hereinafter "Therapeutics"].

An anti-ORFX antibody (*e.g.*, monoclonal antibody) can be used to isolate ORFX by standard techniques, such as affinity chromatography or immunoprecipitation. An anti-ORFX antibody can facilitate the purification of natural ORFX from cells and of recombinantly produced ORFX expressed in host cells. Moreover, an anti-ORFX antibody can be used to detect ORFX protein (*e.g.*, in a cellular lysate or cell supernatant) in order to evaluate the abundance and pattern of expression of the ORFX protein. Anti-ORFX antibodies can be used diagnostically to monitor protein levels in tissue as part of a clinical testing procedure, *e.g.*, to, for example, determine the efficacy of a given treatment regimen. Detection can be facilitated by coupling (*i.e.*, physically linking) the antibody to a detectable substance. Examples of detectable substances include various enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials, and radioactive materials. Examples of suitable enzymes include horseradish peroxidase, alkaline phosphatase, β -galactosidase, or acetylcholinesterase; examples of suitable prosthetic group complexes include streptavidin/biotin and avidin/biotin; examples of suitable fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride or

phycoerythrin; an example of a luminescent material includes luminol; examples of bioluminescent materials include luciferase, luciferin, and aequorin, and examples of suitable radioactive material include ^{125}I , ^{131}I , ^{35}S or ^3H .

ORFX Recombinant Vectors and Host Cells

5 Another aspect of the invention pertains to vectors, preferably expression vectors, containing a nucleic acid encoding ORFX protein, or derivatives, fragments, analogs or homologs thereof. As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of vector is a "plasmid", which refers to a circular double stranded DNA loop into which additional DNA segments can be ligated. Another type of vector is a viral vector, wherein additional DNA segments can be
10 ligated into the viral genome. Certain vectors are capable of autonomous replication in a host cell into which they are introduced (*e.g.*, bacterial vectors having a bacterial origin of replication and episomal mammalian vectors). Other vectors (*e.g.*, non-episomal mammalian vectors) are integrated into the genome of a host cell upon introduction into the host cell, and thereby are replicated along with the host genome. Moreover, certain vectors are capable of directing the expression of genes to which they are operatively linked. Such vectors are referred to herein as "expression vectors". In general, expression vectors of utility in recombinant DNA techniques are often in the form of plasmids. In the present specification, "plasmid" and "vector" can be used interchangeably as the plasmid is the most commonly used form of vector. However, the
15 invention is intended to include such other forms of expression vectors, such as viral vectors (*e.g.*, replication defective retroviruses, adenoviruses and adeno-associated viruses), which serve equivalent functions.

The recombinant expression vectors of the invention comprise a nucleic acid of the invention in a form suitable for expression of the nucleic acid in a host cell, which means that the
25 recombinant expression vectors include one or more regulatory sequences, selected on the basis of the host cells to be used for expression, that is operatively linked to the nucleic acid sequence to be expressed. Within a recombinant expression vector, "operably linked" is intended to mean that the nucleotide sequence of interest is linked to the regulatory sequence(s) in a manner that allows for expression of the nucleotide sequence (*e.g.*, in an *in vitro* transcription/translation system or in a host cell when the vector is introduced into the host cell). The term "regulatory
30 sequence" is intended to include promoters, enhancers and other expression control elements

(e.g., polyadenylation signals). Such regulatory sequences are described, for example, in Goeddel; GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990). Regulatory sequences include those that direct constitutive expression of a nucleotide sequence in many types of host cell and those that direct expression of the nucleotide sequence only in certain host cells (e.g., tissue-specific regulatory sequences). It will be appreciated by those skilled in the art that the design of the expression vector can depend on such factors as the choice of the host cell to be transformed, the level of expression of protein desired, etc. The expression vectors of the invention can be introduced into host cells to thereby produce proteins or peptides, including fusion proteins or peptides, encoded by nucleic acids as described herein (e.g., ORFX proteins, mutant forms of ORFX, fusion proteins, etc.).

The recombinant expression vectors of the invention can be designed for expression of ORFX in prokaryotic or eukaryotic cells. For example, ORFX can be expressed in bacterial cells such as *E. coli*, insect cells (using baculovirus expression vectors) yeast cells or mammalian cells. Suitable host cells are discussed further in Goeddel, GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990). Alternatively, the recombinant expression vector can be transcribed and translated *in vitro*, for example using T7 promoter regulatory sequences and T7 polymerase.

Expression of proteins in prokaryotes is most often carried out in *E. coli* with vectors containing constitutive or inducible promoters directing the expression of either fusion or non-fusion proteins. Fusion vectors add a number of amino acids to a protein encoded therein, usually to the amino terminus of the recombinant protein. Such fusion vectors typically serve three purposes: (1) to increase expression of recombinant protein; (2) to increase the solubility of the recombinant protein; and (3) to aid in the purification of the recombinant protein by acting as a ligand in affinity purification. Often, in fusion expression vectors, a proteolytic cleavage site is introduced at the junction of the fusion moiety and the recombinant protein to enable separation of the recombinant protein from the fusion moiety subsequent to purification of the fusion protein. Such enzymes, and their cognate recognition sequences, include Factor Xa, thrombin and enterokinase. Typical fusion expression vectors include pGEX (Pharmacia Biotech Inc; Smith and Johnson (1988) *Gene* 67:31-40), pMAL (New England Biolabs, Beverly, Mass.) and pRIT5 (Pharmacia, Piscataway, N.J.) that fuse glutathione S-transferase (GST), maltose E binding protein, or protein A, respectively, to the target recombinant protein.

Examples of suitable inducible non-fusion *E. coli* expression vectors include pTrec (Amrann *et al.*, (1988) *Gene* 69:301-315) and pET 11d (Studier *et al.*, GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990) 60-89).

One strategy to maximize recombinant protein expression in *E. coli* is to express the protein in a host bacteria with an impaired capacity to proteolytically cleave the recombinant protein. See, Gottesman, GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990) 119-128. Another strategy is to alter the nucleic acid sequence of the nucleic acid to be inserted into an expression vector so that the individual codons for each amino acid are those preferentially utilized in *E. coli* (Wada *et al.*, (1992) *Nucleic Acids Res.* 20:2111-2118). Such alteration of nucleic acid sequences of the invention can be carried out by standard DNA synthesis techniques.

In another embodiment, the ORFX expression vector is a yeast expression vector. Examples of vectors for expression in yeast *S. cerevisiae* include pYepSec1 (Baldari, *et al.*, (1987) *EMBO J* 6:229-234), pMFa (Kurjan and Herskowitz, (1982) *Cell* 30:933-943), pJRY88 (Schultz *et al.*, (1987) *Gene* 54:113-123), pYES2 (Invitrogen Corporation, San Diego, Calif.), and picZ (Invitrogen Corp, San Diego, Calif.).

Alternatively, ORFX can be expressed in insect cells using baculovirus expression vectors. Baculovirus vectors available for expression of proteins in cultured insect cells (e.g., SF9 cells) include the pAc series (Smith *et al.* (1983) *Mol Cell Biol* 3:2156-2165) and the pVL series (Lucklow and Summers (1989) *Virology* 170:31-39).

In yet another embodiment, a nucleic acid of the invention is expressed in mammalian cells using a mammalian expression vector. Examples of mammalian expression vectors include pCDM8 (Seed (1987) *Nature* 329:840) and pMT2PC (Kaufman *et al.* (1987) *EMBO J* 6: 187-195). When used in mammalian cells, the expression vector's control functions are often provided by viral regulatory elements. For example, commonly used promoters are derived from polyoma, Adenovirus 2, cytomegalovirus and Simian Virus 40. For other suitable expression systems for both prokaryotic and eukaryotic cells. See, e.g., Chapters 16 and 17 of Sambrook *et al.*, MOLECULAR CLONING: A LABORATORY MANUAL, 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989.

In another embodiment, the recombinant mammalian expression vector is capable of directing expression of the nucleic acid preferentially in a particular cell type (e.g., tissue-specific regulatory elements are used to express the nucleic acid). Tissue-specific regulatory elements are known in the art. Non-limiting examples of suitable tissue-specific promoters include the albumin promoter (liver-specific; Pinkert *et al.* (1987) *Genes Dev* 1:268-277), lymphoid-specific promoters (Calame and Eaton (1988) *Adv Immunol* 43:235-275), in particular promoters of T cell receptors (Winoto and Baltimore (1989) *EMBO J* 8:729-733) and immunoglobulins (Banerji *et al.* (1983) *Cell* 33:729-740; Queen and Baltimore (1983) *Cell* 33:741-748), neuron-specific promoters (e.g., the neurofilament promoter; Byrne and Ruddle (1989) *PNAS* 86:5473-5477), pancreas-specific promoters (Edlund *et al.* (1985) *Science* 230:912-916), and mammary gland-specific promoters (e.g., milk whey promoter; U.S. Pat. No. 4,873,316 and European Application Publication No. 264,166). Developmentally-regulated promoters are also encompassed, e.g., the murine hox promoters (Kessel and Gruss (1990) *Science* 249:374-379) and the α -fetoprotein promoter (Campes and Tilghman (1989) *Genes Dev* 3:537-546).

The invention further provides a recombinant expression vector comprising a DNA molecule of the invention cloned into the expression vector in an antisense orientation. That is, the DNA molecule is operatively linked to a regulatory sequence in a manner that allows for expression (by transcription of the DNA molecule) of an RNA molecule that is antisense to ORFX mRNA. Regulatory sequences operatively linked to a nucleic acid cloned in the antisense orientation can be chosen that direct the continuous expression of the antisense RNA molecule in a variety of cell types, for instance viral promoters and/or enhancers, or regulatory sequences can be chosen that direct constitutive, tissue specific or cell type specific expression of antisense RNA. The antisense expression vector can be in the form of a recombinant plasmid, phagemid or attenuated virus in which antisense nucleic acids are produced under the control of a high efficiency regulatory region, the activity of which can be determined by the cell type into which the vector is introduced. For a discussion of the regulation of gene expression using antisense genes see Weintraub *et al.*, "Antisense RNA as a molecular tool for genetic analysis," Reviews--Trends in Genetics, Vol. 1(1) 1986.

Another aspect of the invention pertains to host cells into which a recombinant expression vector of the invention has been introduced. The terms "host cell" and "recombinant

host cell" are used interchangeably herein. It is understood that such terms refer not only to the particular subject cell but to the progeny or potential progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein.

A host cell can be any prokaryotic or eukaryotic cell. For example, ORFX protein can be expressed in bacterial cells such as *E. coli*, insect cells, yeast or mammalian cells (such as Chinese hamster ovary cells (CHO) or COS cells). Other suitable host cells are known to those skilled in the art.

Vector DNA can be introduced into prokaryotic or eukaryotic cells via conventional transformation or transfection techniques. As used herein, the terms "transformation" and "transfection" are intended to refer to a variety of art-recognized techniques for introducing foreign nucleic acid (e.g., DNA) into a host cell, including calcium phosphate or calcium chloride co-precipitation, DEAE-dextran-mediated transfection, lipofection, or electroporation. Suitable methods for transforming or transfecting host cells can be found in Sambrook, *et al.* (MOLECULAR CLONING: A LABORATORY MANUAL, 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989), and other laboratory manuals.

For stable transfection of mammalian cells, it is known that, depending upon the expression vector and transfection technique used, only a small fraction of cells may integrate the foreign DNA into their genome. In order to identify and select these integrants, a gene that encodes a selectable marker (e.g., resistance to antibiotics) is generally introduced into the host cells along with the gene of interest. Various selectable markers include those that confer resistance to drugs, such as G418, hygromycin and methotrexate. Nucleic acid encoding a selectable marker can be introduced into a host cell on the same vector as that encoding ORFX or can be introduced on a separate vector. Cells stably transfected with the introduced nucleic acid can be identified by drug selection (e.g., cells that have incorporated the selectable marker gene will survive, while the other cells die).

A host cell of the invention, such as a prokaryotic or eukaryotic host cell in culture, can be used to produce (i.e., express) ORFX protein. Accordingly, the invention further provides methods for producing ORFX protein using the host cells of the invention. In one embodiment,

the method comprises culturing the host cell of invention (into which a recombinant expression vector encoding ORFX has been introduced) in a suitable medium such that ORFX protein is produced. In another embodiment, the method further comprises isolating ORFX from the medium or the host cell.

Transgenic animals

The host cells of the invention can also be used to produce nonhuman transgenic animals. For example, in one embodiment, a host cell of the invention is a fertilized oocyte or an embryonic stem cell into which ORFX-coding sequences have been introduced. Such host cells can then be used to create non-human transgenic animals in which exogenous ORFX sequences have been introduced into their genome or homologous recombinant animals in which endogenous ORFX sequences have been altered. Such animals are useful for studying the function and/or activity of ORFX and for identifying and/or evaluating modulators of ORFX activity. As used herein, a "transgenic animal" is a non-human animal, preferably a mammal, more preferably a rodent such as a rat or mouse, in which one or more of the cells of the animal includes a transgene. Other examples of transgenic animals include non-human primates, sheep, dogs, cows, goats, chickens, amphibians, etc. A transgene is exogenous DNA that is integrated into the genome of a cell from which a transgenic animal develops and that remains in the genome of the mature animal, thereby directing the expression of an encoded gene product in one or more cell types or tissues of the transgenic animal. As used herein, a "homologous recombinant animal" is a non-human animal, preferably a mammal, more preferably a mouse, in which an endogenous ORFX gene has been altered by homologous recombination between the endogenous gene and an exogenous DNA molecule introduced into a cell of the animal, e.g., an embryonic cell of the animal, prior to development of the animal.

A transgenic animal of the invention can be created by introducing ORFX-encoding nucleic acid into the male pronuclei of a fertilized oocyte, e.g., by microinjection, retroviral infection, and allowing the oocyte to develop in a pseudopregnant female foster animal. The human ORFX DNA sequence of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161) can be introduced as a transgene into the genome of a non-human animal. Alternatively, a nonhuman homologue of the human ORFX gene, such as a mouse ORFX gene, can be isolated based on hybridization to the human ORFX cDNA (described further above) and used as a transgene. Intronic sequences and polyadenylation signals can also be included in the transgene to increase the efficiency of

expression of the transgene. A tissue-specific regulatory sequence(s) can be operably linked to the ORFX transgene to direct expression of ORFX protein to particular cells. Methods for generating transgenic animals via embryo manipulation and microinjection, particularly animals such as mice, have become conventional in the art and are described, for example, in U.S. Pat.

Nos. 4,736,866; 4,870,009; and 4,873,191; and Hogan 1986, In: MANIPULATING THE MOUSE EMBRYO, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. Similar methods are used for production of other transgenic animals. A transgenic founder animal can be identified based upon the presence of the ORFX transgene in its genome and/or expression of ORFX mRNA in tissues or cells of the animals. A transgenic founder animal can then be used to breed additional animals carrying the transgene. Moreover, transgenic animals carrying a transgene encoding ORFX can further be bred to other transgenic animals carrying other transgenes.

To create a homologous recombinant animal, a vector is prepared which contains at least a portion of a ORFX gene into which a deletion, addition or substitution has been introduced to thereby alter, e.g., functionally disrupt, the ORFX gene. The ORFX gene can be a human gene (e.g., SEQ ID NO:2n-1 (wherein $n = 1$ to 3161)), but more preferably, is a non-human homologue of a human ORFX gene. For example, a mouse homologue of human ORFX gene of SEQ ID NO:2n-1 (wherein $n = 1$ to 3161) can be used to construct a homologous recombination vector suitable for altering an endogenous ORFX gene in the mouse genome. In one embodiment, the vector is designed such that, upon homologous recombination, the endogenous ORFX gene is functionally disrupted (*i.e.*, no longer encodes a functional protein; also referred to as a "knock out" vector).

Alternatively, the vector can be designed such that, upon homologous recombination, the endogenous ORFX gene is mutated or otherwise altered but still encodes functional protein (e.g., the upstream regulatory region can be altered to thereby alter the expression of the endogenous ORFX protein). In the homologous recombination vector, the altered portion of the ORFX gene is flanked at its 5' and 3' ends by additional nucleic acid of the ORFX gene to allow for homologous recombination to occur between the exogenous ORFX gene carried by the vector and an endogenous ORFX gene in an embryonic stem cell. The additional flanking ORFX nucleic acid is of sufficient length for successful homologous recombination with the endogenous gene. Typically, several kilobases of flanking DNA (both at the 5' and 3' ends) are included in the vector. See e.g., Thomas *et al.* (1987) *Cell* 51:503 for a description of

homologous recombination vectors. The vector is introduced into an embryonic stem cell line (e.g., by electroporation) and cells in which the introduced ORFX gene has homologously recombined with the endogenous ORFX gene are selected (see e.g., Li *et al.* (1992) *Cell* 69:915).

The selected cells are then injected into a blastocyst of an animal (e.g., a mouse) to form aggregation chimeras. See e.g., Bradley 1987, In: TERATOCARCINOMAS AND EMBRYONIC STEM CELLS: A PRACTICAL APPROACH, Robertson, ed. IRL, Oxford, pp. 113-152. A chimeric embryo can then be implanted into a suitable pseudopregnant female foster animal and the embryo brought to term. Progeny harboring the homologously recombined DNA in their germ cells can be used to breed animals in which all cells of the animal contain the homologously recombined DNA by germline transmission of the transgene. Methods for constructing homologous recombination vectors and homologous recombinant animals are described further in Bradley (1991) *Curr Opin Biotechnol* 2:823-829; PCT International Publication Nos.: WO 90/11354; WO 91/01140; WO 92/0968; and WO 93/04169.

In another embodiment, transgenic non-humans animals can be produced that contain selected systems that allow for regulated expression of the transgene. One example of such a system is the cre/loxP recombinase system of bacteriophage P1. For a description of the cre/loxP recombinase system, see, e.g., Lakso *et al.* (1992) *PNAS* 89:6232-6236. Another example of a recombinase system is the FLP recombinase system of *Saccharomyces cerevisiae* (O'Gorman *et al.* (1991) *Science* 251:1351-1355. If a cre/loxP recombinase system is used to regulate expression of the transgene, animals containing transgenes encoding both the Cre recombinase and a selected protein are required. Such animals can be provided through the construction of "double" transgenic animals, e.g., by mating two transgenic animals, one containing a transgene encoding a selected protein and the other containing a transgene encoding a recombinase.

Clones of the non-human transgenic animals described herein can also be produced according to the methods described in Wilmot *et al.* (1997) *Nature* 385:810-813. In brief, a cell, e.g., a somatic cell, from the transgenic animal can be isolated and induced to exit the growth cycle and enter G₀ phase. The quiescent cell can then be fused, e.g., through the use of electrical pulses, to an enucleated oocyte from an animal of the same species from which the quiescent cell is isolated. The reconstructed oocyte is then cultured such that it develops to morula or blastocyte and then transferred to pseudopregnant female foster animal. The offspring borne of

this female foster animal will be a clone of the animal from which the cell, *e.g.*, the somatic cell, is isolated.

Pharmaceutical Compositions

The ORFX nucleic acid molecules, ORFX proteins, and anti-ORFX antibodies (also referred to herein as "active compounds") of the invention, and derivatives, fragments, analogs and homologs thereof, can be incorporated into pharmaceutical compositions suitable for administration. Such compositions typically comprise the nucleic acid molecule, protein, or antibody and a pharmaceutically acceptable carrier. As used herein, "pharmaceutically acceptable carrier" is intended to include any and all solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents, and the like, compatible with pharmaceutical administration. Suitable carriers are described in the most recent edition of Remington's Pharmaceutical Sciences, a standard reference text in the field, which is incorporated herein by reference. Preferred examples of such carriers or diluents include, but are not limited to, water, saline, finger's solutions, dextrose solution, and 5% human serum albumin. Liposomes and non-aqueous vehicles such as fixed oils may also be used. The use of such media and agents for pharmaceutically active substances is well known in the art. Except insofar as any conventional media or agent is incompatible with the active compound, use thereof in the compositions is contemplated. Supplementary active compounds can also be incorporated into the compositions.

A pharmaceutical composition of the invention is formulated to be compatible with its intended route of administration. Examples of routes of administration include parenteral, *e.g.*, intravenous, intradermal, subcutaneous, oral (*e.g.*, inhalation), transdermal (topical), transmucosal, and rectal administration. Solutions or suspensions used for parenteral, intradermal, or subcutaneous application can include the following components: a sterile diluent such as water for injection, saline solution, fixed oils, polyethylene glycols, glycerine, propylene glycol or other synthetic solvents; antibacterial agents such as benzyl alcohol or methyl parabens; antioxidants such as ascorbic acid or sodium bisulfite; chelating agents such as ethylenediaminetetraacetic acid; buffers such as acetates, citrates or phosphates, and agents for the adjustment of tonicity such as sodium chloride or dextrose. The pH can be adjusted with acids or bases, such as hydrochloric acid or sodium hydroxide. The parenteral preparation can be enclosed in ampoules, disposable syringes or multiple dose vials made of glass or plastic.

Pharmaceutical compositions suitable for injectable use include sterile aqueous solutions (where water soluble) or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersion. For intravenous administration, suitable carriers include physiological saline, bacteriostatic water, Cremophor EL™ (BASF, Parsippany, N.J.) or phosphate buffered saline (PBS). In all cases, the composition must be sterile and should be fluid to the extent that easy syringeability exists. It must be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol, and liquid polyethylene glycol, and the like), and suitable mixtures thereof. The proper fluidity can be maintained, for example, by the use of a coating such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. Prevention of the action of microorganisms can be achieved by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, ascorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars, polyalcohols such as manitol, sorbitol, sodium chloride in the composition. Prolonged absorption of the injectable compositions can be brought about by including in the composition an agent which delays absorption, for example, aluminum monostearate and gelatin.

Sterile injectable solutions can be prepared by incorporating the active compound (*e.g.*, a ORFX protein or anti-ORFX antibody) in the required amount in an appropriate solvent with one or a combination of ingredients enumerated above, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the active compound into a sterile vehicle that contains a basic dispersion medium and the required other ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, methods of preparation are vacuum drying and freeze-drying that yields a powder of the active ingredient plus any additional desired ingredient from a previously sterile-filtered solution thereof.

Oral compositions generally include an inert diluent or an edible carrier. They can be enclosed in gelatin capsules or compressed into tablets. For the purpose of oral therapeutic administration, the active compound can be incorporated with excipients and used in the form of tablets, troches, or capsules. Oral compositions can also be prepared using a fluid carrier for use

as a mouthwash, wherein the compound in the fluid carrier is applied orally and swished and expectorated or swallowed. Pharmaceutically compatible binding agents, and/or adjuvant materials can be included as part of the composition. The tablets, pills, capsules, troches and the like can contain any of the following ingredients, or compounds of a similar nature: a binder such as microcrystalline cellulose, gum tragacanth or gelatin; an excipient such as starch or lactose, a disintegrating agent such as alginic acid, Primogel, or corn starch; a lubricant such as magnesium stearate or Sterotes; a glidant such as colloidal silicon dioxide; a sweetening agent such as sucrose or saccharin; or a flavoring agent such as peppermint, methyl salicylate, or orange flavoring.

For administration by inhalation, the compounds are delivered in the form of an aerosol spray from pressured container or dispenser which contains a suitable propellant, *e.g.*, a gas such as carbon dioxide, or a nebulizer.

Systemic administration can also be by transmucosal or transdermal means. For transmucosal or transdermal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art, and include, for example, for transmucosal administration, detergents, bile salts, and fusidic acid derivatives. Transmucosal administration can be accomplished through the use of nasal sprays or suppositories. For transdermal administration, the active compounds are formulated into ointments, salves, gels, or creams as generally known in the art.

The compounds can also be prepared in the form of suppositories (*e.g.*, with conventional suppository bases such as cocoa butter and other glycerides) or retention enemas for rectal delivery.

In one embodiment, the active compounds are prepared with carriers that will protect the compound against rapid elimination from the body, such as a controlled release formulation, including implants and microencapsulated delivery systems. Biodegradable, biocompatible polymers can be used, such as ethylene vinyl acetate, polyanhydrides, polyglycolic acid, collagen, polyorthoesters, and polylactic acid. Methods for preparation of such formulations will be apparent to those skilled in the art. The materials can also be obtained commercially from Alza Corporation and Nova Pharmaceuticals, Inc. Liposomal suspensions (including liposomes targeted to infected cells with monoclonal antibodies to viral antigens) can also be used as pharmaceutically acceptable carriers. These can be prepared according to methods known to those skilled in the art, for example, as described in U.S. Pat. No. 4,522,811.

It is especially advantageous to formulate oral or parenteral compositions in dosage unit form for ease of administration and uniformity of dosage. Dosage unit form as used herein refers to physically discrete units suited as unitary dosages for the subject to be treated; each unit containing a predetermined quantity of active compound calculated to produce the desired therapeutic effect in association with the required pharmaceutical carrier. The specification for the dosage unit forms of the invention are dictated by and directly dependent on the unique characteristics of the active compound and the particular therapeutic effect to be achieved.

The nucleic acid molecules of the invention can be inserted into vectors and used as gene therapy vectors. Gene therapy vectors can be delivered to a subject by any of a number of routes, *e.g.*, as described in U.S. Patent Nos. 5,703,055. Delivery can thus also include, *e.g.*, intravenous injection, local administration (see U.S. Pat. No. 5,328,470) or stereotactic injection (see *e.g.*, Chen *et al.* (1994) *PNAS* 91:3054-3057). The pharmaceutical preparation of the gene therapy vector can include the gene therapy vector in an acceptable diluent, or can comprise a slow release matrix in which the gene delivery vehicle is imbedded. Alternatively, where the complete gene delivery vector can be produced intact from recombinant cells, *e.g.*, retroviral vectors, the pharmaceutical preparation can include one or more cells that produce the gene delivery system.

The pharmaceutical compositions can be included in a container, pack, or dispenser together with instructions for administration.

Additional Uses and Methods of the Invention

The nucleic acid molecules, proteins, protein homologues, and antibodies described herein can be used in one or more of the following methods: (a) screening assays; (b) detection assays (*e.g.*, chromosomal mapping, cell and tissue typing, forensic biology), (c) predictive medicine (*e.g.*, diagnostic assays, prognostic assays, monitoring clinical trials, and pharmacogenomics); and (d) methods of treatment (*e.g.*, therapeutic and prophylactic).

The isolated nucleic acid molecules of the invention can be used to express ORFX protein (*e.g.*, via a recombinant expression vector in a host cell in gene therapy applications), to detect ORFX mRNA (*e.g.*, in a biological sample) or a genetic lesion in a ORFX gene, and to modulate ORFX activity, as described further below. In addition, the ORFX proteins can be used to screen drugs or compounds that modulate the ORFX activity or expression as well as to treat disorders characterized by insufficient or excessive production of ORFX protein, for

example proliferative or differentiative disorders, or production of ORFX protein forms that have decreased or aberrant activity compared to ORFX wild type protein. In addition, the anti-ORFX antibodies of the invention can be used to detect and isolate ORFX proteins and modulate ORFX activity.

This invention further pertains to novel agents identified by the above described screening assays and uses thereof for treatments as described herein.

Screening Assays

The invention provides a method (also referred to herein as a "screening assay") for identifying modulators, *i.e.*, candidate or test compounds or agents (*e.g.*, peptides, peptidomimetics, small molecules or other drugs) that bind to ORFX proteins or have a stimulatory or inhibitory effect on, for example, ORFX expression or ORFX activity.

In one embodiment, the invention provides assays for screening candidate or test compounds which bind to or modulate the activity of a ORFX protein or polypeptide or biologically active portion thereof. The test compounds of the present invention can be obtained using any of the numerous approaches in combinatorial library methods known in the art, including: biological libraries; spatially addressable parallel solid phase or solution phase libraries; synthetic library methods requiring deconvolution; the "one-bead one-compound" library method; and synthetic library methods using affinity chromatography selection. The biological library approach is limited to peptide libraries, while the other four approaches are applicable to peptide, non-peptide oligomer or small molecule libraries of compounds (Lam (1997) *Anticancer Drug Des* 12:145).

Examples of methods for the synthesis of molecular libraries can be found in the art, for example in: DeWitt *et al.* (1993) *Proc Natl Acad Sci U.S.A.* 90:6909; Erb *et al.* (1994) *Proc Natl Acad Sci U.S.A.* 91:11422; Zuckermann *et al.* (1994) *J Med Chem* 37:2678; Cho *et al.* (1993) *Science* 261:1303; Carrell *et al.* (1994) *Angew Chem Int Ed Engl* 33:2059; Carell *et al.* (1994) *Angew Chem Int Ed Engl* 33:2061; and Gallop *et al.* (1994) *J Med Chem* 37:1233.

Libraries of compounds may be presented in solution (*e.g.*, Houghten (1992) *Biotechniques* 13:412-421), or on beads (Lam (1991) *Nature* 354:82-84), on chips (Fodor (1993) *Nature* 364:555-556), bacteria (Ladner U.S. Pat. No. 5,223,409), spores (Ladner USP '409), plasmids (Cull *et al.* (1992) *Proc Natl Acad Sci USA* 89:1865-1869) or on phage (Scott and

Smith (1990) *Science* 249:386-390; Devlin (1990) *Science* 249:404-406; Cwirla *et al.* (1990) *Proc Natl Acad Sci U.S.A.* 87:6378-6382; Felici (1991) *J Mol Biol* 222:301-310; Ladner *above.*).

In one embodiment, an assay is a cell-based assay in which a cell which expresses a membrane-bound form of ORFX protein, or a biologically active portion thereof, on the cell surface is contacted with a test compound and the ability of the test compound to bind to a ORFX protein determined. The cell, for example, can be of mammalian origin or a yeast cell. Determining the ability of the test compound to bind to the ORFX protein can be accomplished, for example, by coupling the test compound with a radioisotope or enzymatic label such that binding of the test compound to the ORFX protein or biologically active portion thereof can be determined by detecting the labeled compound in a complex. For example, test compounds can be labeled with ^{125}I , ^{35}S , ^{14}C , or ^3H , either directly or indirectly, and the radioisotope detected by direct counting of radioemission or by scintillation counting. Alternatively, test compounds can be enzymatically labeled with, for example, horseradish peroxidase, alkaline phosphatase, or luciferase, and the enzymatic label detected by determination of conversion of an appropriate substrate to product. In one embodiment, the assay comprises contacting a cell which expresses a membrane-bound form of ORFX protein, or a biologically active portion thereof, on the cell surface with a known compound which binds ORFX to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with a ORFX protein, wherein determining the ability of the test compound to interact with a ORFX protein comprises determining the ability of the test compound to preferentially bind to ORFX or a biologically active portion thereof as compared to the known compound.

In another embodiment, an assay is a cell-based assay comprising contacting a cell expressing a membrane-bound form of ORFX protein, or a biologically active portion thereof, on the cell surface with a test compound and determining the ability of the test compound to modulate (*e.g.*, stimulate or inhibit) the activity of the ORFX protein or biologically active portion thereof. Determining the ability of the test compound to modulate the activity of ORFX or a biologically active portion thereof can be accomplished, for example, by determining the ability of the ORFX protein to bind to or interact with a ORFX target molecule. As used herein, a "target molecule" is a molecule with which a ORFX protein binds or interacts in nature, for example, a molecule on the surface of a cell which expresses a ORFX interacting protein, a molecule on the surface of a second cell, a molecule in the extracellular milieu, a molecule

associated with the internal surface of a cell membrane or a cytoplasmic molecule. A ORFX target molecule can be a non-ORFX molecule or a ORFX protein or polypeptide of the present invention. In one embodiment, a ORFX target molecule is a component of a signal transduction pathway that facilitates transduction of an extracellular signal (e.g., a signal generated by binding of a compound to a membrane-bound ORFX molecule) through the cell membrane and into the cell. The target, for example, can be a second intercellular protein that has catalytic activity or a protein that facilitates the association of downstream signaling molecules with ORFX.

Determining the ability of the ORFX protein to bind to or interact with a ORFX target molecule can be accomplished by one of the methods described above for determining direct binding. In one embodiment, determining the ability of the ORFX protein to bind to or interact with a ORFX target molecule can be accomplished by determining the activity of the target molecule. For example, the activity of the target molecule can be determined by detecting induction of a cellular second messenger of the target (*i.e.* intracellular Ca^{2+} , diacylglycerol, IP_3 , etc.), detecting catalytic/enzymatic activity of the target on an appropriate substrate, detecting induction of a reporter gene (comprising a ORFX-responsive regulatory element operatively linked to a nucleic acid encoding a detectable marker, *e.g.*, luciferase), or detecting a cellular response, for example, cell survival, cellular differentiation, or cell proliferation.

In yet another embodiment, an assay of the present invention is a cell-free assay comprising contacting a ORFX protein or biologically active portion thereof with a test compound and determining the ability of the test compound to bind to the ORFX protein or biologically active portion thereof. Binding of the test compound to the ORFX protein can be determined either directly or indirectly as described above. In one embodiment, the assay comprises contacting the ORFX protein or biologically active portion thereof with a known compound which binds ORFX to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with a ORFX protein, wherein determining the ability of the test compound to interact with a ORFX protein comprises determining the ability of the test compound to preferentially bind to ORFX or biologically active portion thereof as compared to the known compound.

In another embodiment, an assay is a cell-free assay comprising contacting ORFX protein or biologically active portion thereof with a test compound and determining the ability of the test compound to modulate (*e.g.*, stimulate or inhibit) the activity of the ORFX protein or

biologically active portion thereof. Determining the ability of the test compound to modulate the activity of ORFX can be accomplished, for example, by determining the ability of the ORFX protein to bind to a ORFX target molecule by one of the methods described above for determining direct binding. In an alternative embodiment, determining the ability of the test compound to modulate the activity of ORFX can be accomplished by determining the ability of the ORFX protein further modulate a ORFX target molecule. For example, the catalytic/enzymatic activity of the target molecule on an appropriate substrate can be determined as previously described.

In yet another embodiment, the cell-free assay comprises contacting the ORFX protein or biologically active portion thereof with a known compound which binds ORFX to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with a ORFX protein, wherein determining the ability of the test compound to interact with a ORFX protein comprises determining the ability of the ORFX protein to preferentially bind to or modulate the activity of a ORFX target molecule.

The cell-free assays of the present invention are amenable to use of both the soluble form or the membrane-bound form of ORFX. In the case of cell-free assays comprising the membrane-bound form of ORFX, it may be desirable to utilize a solubilizing agent such that the membrane-bound form of ORFX is maintained in solution. Examples of such solubilizing agents include non-ionic detergents such as n-octylglucoside, n-dodecylglucoside, n-dodecylmaltoide, octanoyl-N-methylglucamide, decanoyl-N-methylglucamide, Triton® X-100, Triton® X-114, Thesit®, Isotridecypoly(ethylene glycol ether)_m, N-dodecyl-N,N-dimethyl-3-ammonio-1-propane sulfonate, 3-(3-cholamidopropyl)dimethylamminiol-1-propane sulfonate (CHAPS), or 3-(3-cholamidopropyl)dimethylamminiol-2-hydroxy-1-propane sulfonate (CHAPSO).

In more than one embodiment of the above assay methods of the present invention, it may be desirable to immobilize either ORFX or its target molecule to facilitate separation of complexed from uncomplexed forms of one or both of the proteins, as well as to accommodate automation of the assay. Binding of a test compound to ORFX, or interaction of ORFX with a target molecule in the presence and absence of a candidate compound, can be accomplished in any vessel suitable for containing the reactants. Examples of such vessels include microtiter plates, test tubes, and micro-centrifuge tubes. In one embodiment, a fusion protein can be provided that adds a domain that allows one or both of the proteins to be bound to a matrix. For

example, GST-ORFX fusion proteins or GST-target fusion proteins can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, MO) or glutathione derivatized microtiter plates, that are then combined with the test compound or the test compound and either the non-adsorbed target protein or ORFX protein, and the mixture is incubated under conditions conducive to complex formation (*e.g.*, at physiological conditions for salt and pH). Following incubation, the beads or microtiter plate wells are washed to remove any unbound components, the matrix immobilized in the case of beads, complex determined either directly or indirectly, for example, as described above. Alternatively, the complexes can be dissociated from the matrix, and the level of ORFX binding or activity determined using standard techniques.

Other techniques for immobilizing proteins on matrices can also be used in the screening assays of the invention. For example, either ORFX or its target molecule can be immobilized utilizing conjugation of biotin and streptavidin. Biotinylated ORFX or target molecules can be prepared from biotin-NHS (N-hydroxy-succinimide) using techniques well known in the art (*e.g.*, biotinylation kit, Pierce Chemicals, Rockford, Ill.), and immobilized in the wells of streptavidin-coated 96 well plates (Pierce Chemical). Alternatively, antibodies reactive with ORFX or target molecules, but which do not interfere with binding of the ORFX protein to its target molecule, can be derivatized to the wells of the plate, and unbound target or ORFX trapped in the wells by antibody conjugation. Methods for detecting such complexes, in addition to those described above for the GST-immobilized complexes, include immunodetection of complexes using antibodies reactive with the ORFX or target molecule, as well as enzyme-linked assays that rely on detecting an enzymatic activity associated with the ORFX or target molecule.

In another embodiment, modulators of ORFX expression are identified in a method wherein a cell is contacted with a candidate compound and the expression of ORFX mRNA or protein in the cell is determined. The level of expression of ORFX mRNA or protein in the presence of the candidate compound is compared to the level of expression of ORFX mRNA or protein in the absence of the candidate compound. The candidate compound can then be identified as a modulator of ORFX expression based on this comparison. For example, when expression of ORFX mRNA or protein is greater (statistically significantly greater) in the presence of the candidate compound than in its absence, the candidate compound is identified as a stimulator of ORFX mRNA or protein expression. Alternatively, when expression of ORFX mRNA or protein is less (statistically significantly less) in the presence of the candidate

compound than in its absence, the candidate compound is identified as an inhibitor of ORFX mRNA or protein expression. The level of ORFX mRNA or protein expression in the cells can be determined by methods described herein for detecting ORFX mRNA or protein.

In yet another aspect of the invention, the ORFX proteins can be used as "bait proteins" in a two-hybrid assay or three hybrid assay (see, *e.g.*, U.S. Pat. No. 5,283,317; Zervos *et al.* (1993) Cell 72:223-232; Madura *et al.* (1993) J Biol Chem 268:12046-12054; Bartel *et al.* (1993) Biotechniques 14:920-924; Iwabuchi *et al.* (1993) Oncogene 8:1693-1696; and Brent WO94/10300), to identify other proteins that bind to or interact with ORFX ("ORFX-binding proteins" or "ORFX-bp") and modulate ORFX activity. Such ORFX-binding proteins are also likely to be involved in the propagation of signals by the ORFX proteins as, for example, upstream or downstream elements of the ORFX pathway.

The two-hybrid system is based on the modular nature of most transcription factors, which consist of separable DNA-binding and activation domains. Briefly, the assay utilizes two different DNA constructs. In one construct, the gene that codes for ORFX is fused to a gene encoding the DNA binding domain of a known transcription factor (*e.g.*, GAL-4). In the other construct, a DNA sequence, from a library of DNA sequences, that encodes an unidentified protein ("prey" or "sample") is fused to a gene that codes for the activation domain of the known transcription factor. If the "bait" and the "prey" proteins are able to interact, *in vivo*, forming a ORFX-dependent complex, the DNA-binding and activation domains of the transcription factor are brought into close proximity. This proximity allows transcription of a reporter gene (*e.g.*, LacZ) that is operably linked to a transcriptional regulatory site responsive to the transcription factor. Expression of the reporter gene can be detected and cell colonies containing the functional transcription factor can be isolated and used to obtain the cloned gene that encodes the protein which interacts with ORFX.

This invention further pertains to novel agents identified by the above-described screening assays and uses thereof for treatments as described herein.

Detection Assays

Portions or fragments of the cDNA sequences identified herein (and the corresponding complete gene sequences) can be used in numerous ways as polynucleotide reagents. For example, these sequences can be used to: (i) map their respective genes on a chromosome; and, thus, locate gene regions associated with genetic disease; (ii) identify an individual from a

minute biological sample (tissue typing); and (iii) aid in forensic identification of a biological sample.

The ORFX sequences of the present invention can also be used to identify individuals from minute biological samples. In this technique, an individual's genomic DNA is digested with one or more restriction enzymes, and probed on a Southern blot to yield unique bands for identification. The sequences of the present invention are useful as additional DNA markers for RFLP ("restriction fragment length polymorphisms," described in U.S. Pat. No. 5,272,057).

Furthermore, the sequences of the present invention can be used to provide an alternative technique that determines the actual base-by-base DNA sequence of selected portions of an individual's genome. Thus, the ORFX sequences described herein can be used to prepare two PCR primers from the 5' and 3' ends of the sequences. These primers can then be used to amplify an individual's DNA and subsequently sequence it.

Panels of corresponding DNA sequences from individuals, prepared in this manner, can provide unique individual identifications, as each individual will have a unique set of such DNA sequences due to allelic differences. The sequences of the present invention can be used to obtain such identification sequences from individuals and from tissue. The ORFX sequences of the invention uniquely represent portions of the human genome. Allelic variation occurs to some degree in the coding regions of these sequences, and to a greater degree in the noncoding regions. It is estimated that allelic variation between individual humans occurs with a frequency of about once per each 500 bases. Much of the allelic variation is due to single nucleotide polymorphisms (SNPs), which include restriction fragment length polymorphisms (RFLPs).

Each of the sequences described herein can, to some degree, be used as a standard against which DNA from an individual can be compared for identification purposes. Because greater numbers of polymorphisms occur in the noncoding regions, fewer sequences are necessary to differentiate individuals. The noncoding sequences of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161), as described above, can comfortably provide positive individual identification with a panel of perhaps 10 to 1,000 primers that each yield a noncoding amplified sequence of 100 bases. If predicted coding sequences are used, a more appropriate number of primers for positive individual identification would be 500-2,000.

Predictive Medicine

The present invention also pertains to the field of predictive medicine in which diagnostic assays, prognostic assays, pharmacogenomics, and monitoring clinical trials are used for prognostic (predictive) purposes to thereby treat an individual prophylactically. Accordingly, one aspect of the present invention relates to diagnostic assays for determining ORFX protein and/or nucleic acid expression as well as ORFX activity, in the context of a biological sample (e.g., blood, serum, cells, tissue) to thereby determine whether an individual is afflicted with a disease or disorder, or is at risk of developing a disorder, associated with aberrant ORFX expression or activity. The invention also provides for prognostic (or predictive) assays for determining whether an individual is at risk of developing a disorder associated with ORFX protein, nucleic acid expression or activity. For example, mutations in a ORFX gene can be assayed in a biological sample. Such assays can be used for prognostic or predictive purpose to thereby prophylactically treat an individual prior to the onset of a disorder characterized by or associated with ORFX protein, nucleic acid expression or activity.

Another aspect of the invention provides methods for determining ORFX protein, nucleic acid expression or ORFX activity in an individual to thereby select appropriate therapeutic or prophylactic agents for that individual (referred to herein as "pharmacogenomics"). Pharmacogenomics allows for the selection of agents (e.g., drugs) for therapeutic or prophylactic treatment of an individual based on the genotype of the individual (e.g., the genotype of the individual examined to determine the ability of the individual to respond to a particular agent.)

Yet another aspect of the invention pertains to monitoring the influence of agents (e.g., drugs, compounds) on the expression or activity of ORFX in clinical trials.

Use of Partial ORFX Sequences in Forensic Biology

DNA-based identification techniques can also be used in forensic biology. Forensic biology is a scientific field employing genetic typing of biological evidence found at a crime scene as a means for positively identifying, for example, a perpetrator of a crime. To make such an identification, PCR technology can be used to amplify DNA sequences taken from very small biological samples such as tissues, e.g., hair or skin, or body fluids, e.g., blood, saliva, or semen

found at a crime scene. The amplified sequence can then be compared to a standard, thereby allowing identification of the origin of the biological sample.

The sequences of the present invention can be used to provide polynucleotide reagents, *e.g.*, PCR primers, targeted to specific loci in the human genome, that can enhance the reliability of DNA-based forensic identifications by, for example, providing another "identification marker" (*i.e.* another DNA sequence that is unique to a particular individual). As mentioned above, actual base sequence information can be used for identification as an accurate alternative to patterns formed by restriction enzyme generated fragments. Sequences targeted to noncoding regions of SEQ ID NOs: __ are particularly appropriate for this use as greater numbers of polymorphisms occur in the noncoding regions, making it easier to differentiate individuals using this technique. Examples of polynucleotide reagents include the ORFX sequences or portions thereof, *e.g.*, fragments derived from the noncoding regions of one or more of SEQ ID NO:2*n*-1 (where $n = 1$ to 3161), having a length of at least 20 bases, preferably at least 30 bases.

The ORFX sequences described herein can further be used to provide polynucleotide reagents, *e.g.*, labeled or label-able probes that can be used, for example, in an *in situ* hybridization technique, to identify a specific tissue, *e.g.*, brain tissue, etc. This can be very useful in cases where a forensic pathologist is presented with a tissue of unknown origin. Panels of such ORFX probes can be used to identify tissue by species and/or by organ type.

In a similar fashion, these reagents, *e.g.*, ORFX primers or probes can be used to screen tissue culture for contamination (*i.e.* screen for the presence of a mixture of different types of cells in a culture).

Predictive Medicine

The present invention also pertains to the field of predictive medicine in which diagnostic assays, prognostic assays, pharmacogenomics, and monitoring clinical trials are used for prognostic (predictive) purposes to thereby treat an individual prophylactically. Accordingly, one aspect of the present invention relates to diagnostic assays for determining ORFX protein and/or nucleic acid expression as well as ORFX activity, in the context of a biological sample (*e.g.*, blood, serum, cells, tissue) to thereby determine whether an individual is afflicted with a disease or disorder, or is at risk of developing a disorder, associated with aberrant ORFX expression or activity. The invention also provides for prognostic (or predictive) assays for determining whether an individual is at risk of developing a disorder associated with ORFX protein, nucleic

acid expression or activity. For example, mutations in a ORFX gene can be assayed in a biological sample. Such assays can be used for prognostic or predictive purpose to thereby prophylactically treat an individual prior to the onset of a disorder characterized by or associated with ORFX protein, nucleic acid expression or activity.

Another aspect of the invention provides methods for determining ORFX protein, nucleic acid expression or ORFX activity in an individual to thereby select appropriate therapeutic or prophylactic agents for that individual (referred to herein as "pharmacogenomics"). Pharmacogenomics allows for the selection of agents (*e.g.*, drugs) for therapeutic or prophylactic treatment of an individual based on the genotype of the individual (*e.g.*, the genotype of the individual examined to determine the ability of the individual to respond to a particular agent.)

Yet another aspect of the invention pertains to monitoring the influence of agents (*e.g.*, drugs, compounds) on the expression or activity of ORFX in clinical trials.

These and other agents are described in further detail in the following sections.

Diagnostic Assays

Other conditions in which proliferation of cells plays a role include tumors, restenosis, psoriasis, Dupuytren's contracture, diabetic complications, Kaposi's sarcoma and rheumatoid arthritis.

An ORFX polypeptide may be used to identify an interacting polypeptide a sample or tissue. The method comprises contacting the sample or tissue with ORFX, allowing formation of a complex between the ORFX polypeptide and the interacting polypeptide, and detecting the complex, if present.

The proteins of the invention may be used to stimulate production of antibodies specifically binding the proteins. Such antibodies may be used in immunodiagnostic procedures to detect the occurrence of the protein in a sample. The proteins of the invention may be used to stimulate cell growth and cell proliferation in conditions in which such growth would be favorable. An example would be to counteract toxic side effects of chemotherapeutic agents on, for example, hematopoiesis and platelet formation, linings of the gastrointestinal tract, and hair follicles. They may also be used to stimulate new cell growth in neurological disorders including, for example, Alzheimer's disease. Alternatively, antagonistic treatments may be administered in which an antibody specifically binding the ORFX-like proteins of the invention

would abrogate the specific growth-inducing effects of the proteins. Such antibodies may be useful, for example, in the treatment of proliferative disorders including various tumors and benign hyperplasias.

Polynucleotides or oligonucleotides corresponding to any one portion of the ORFX nucleic acids of SEQ ID NO:2*n*-1 (wherein $n = 1$ to 3161) may be used to detect DNA containing a corresponding ORF gene, or detect the expression of a corresponding ORFX gene, or ORFX-like gene. For example, an ORFX nucleic acid expressed in a particular cell or tissue, as noted in Table 2, can be used to identify the presence of that particular cell type.

An exemplary method for detecting the presence or absence of ORFX in a biological sample involves obtaining a biological sample from a test subject and contacting the biological sample with a compound or an agent capable of detecting ORFX protein or nucleic acid (*e.g.*, mRNA, genomic DNA) that encodes ORFX protein such that the presence of ORFX is detected in the biological sample. An agent for detecting ORFX mRNA or genomic DNA is a labeled nucleic acid probe capable of hybridizing to ORFX mRNA or genomic DNA. The nucleic acid probe can be, for example, a full-length ORFX nucleic acid, such as the nucleic acid of SEQ ID NO:2*n*-1 (wherein $n = 1$ to 3161), or a portion thereof, such as an oligonucleotide of at least 15, 30, 50, 100, 250 or 500 nucleotides in length and sufficient to specifically hybridize under stringent conditions to ORFX mRNA or genomic DNA, as described above. Other suitable probes for use in the diagnostic assays of the invention are described herein.

An agent for detecting ORFX protein is an antibody capable of binding to ORFX protein, preferably an antibody with a detectable label. Antibodies can be polyclonal, or more preferably, monoclonal. An intact antibody, or a fragment thereof (*e.g.*, Fab or F(ab')₂) can be used. The term "labeled", with regard to the probe or antibody, is intended to encompass direct labeling of the probe or antibody by coupling (*i.e.*, physically linking) a detectable substance to the probe or antibody, as well as indirect labeling of the probe or antibody by reactivity with another reagent that is directly labeled. Examples of indirect labeling include detection of a primary antibody using a fluorescently labeled secondary antibody and end-labeling of a DNA probe with biotin such that it can be detected with fluorescently labeled streptavidin. The term "biological sample" is intended to include tissues, cells and biological fluids isolated from a subject, as well as tissues, cells and fluids present within a subject. That is, the detection method of the invention can be used to detect ORFX mRNA, protein, or genomic DNA in a biological sample *in vitro* as

well as *in vivo*. For example, *in vitro* techniques for detection of ORFX mRNA include Northern hybridizations and *in situ* hybridizations. *In vitro* techniques for detection of ORFX protein include enzyme linked immunosorbent assays (ELISAs), Western blots, immunoprecipitations and immunofluorescence. *In vitro* techniques for detection of ORFX genomic DNA include Southern hybridizations. Furthermore, *in vivo* techniques for detection of ORFX protein include introducing into a subject a labeled anti-ORFX antibody. For example, the antibody can be labeled with a radioactive marker whose presence and location in a subject can be detected by standard imaging techniques.

In one embodiment, the biological sample contains protein molecules from the test subject. Alternatively, the biological sample can contain mRNA molecules from the test subject or genomic DNA molecules from the test subject. A preferred biological sample is a peripheral blood leukocyte sample isolated by conventional means from a subject.

In another embodiment, the methods further involve obtaining a control biological sample from a control subject, contacting the control sample with a compound or agent capable of detecting ORFX protein, mRNA, or genomic DNA, such that the presence of ORFX protein, mRNA or genomic DNA is detected in the biological sample, and comparing the presence of ORFX protein, mRNA or genomic DNA in the control sample with the presence of ORFX protein, mRNA or genomic DNA in the test sample.

The invention also encompasses kits for detecting the presence of ORFX in a biological sample. For example, the kit can comprise: a labeled compound or agent capable of detecting ORFX protein or mRNA in a biological sample; means for determining the amount of ORFX in the sample; and means for comparing the amount of ORFX in the sample with a standard. The compound or agent can be packaged in a suitable container. The kit can further comprise instructions for using the kit to detect ORFX protein or nucleic acid.

Prognostic Assays

The diagnostic methods described herein can furthermore be utilized to identify subjects having or at risk of developing a disease or disorder associated with aberrant ORFX expression or activity. For example, the assays described herein, such as the preceding diagnostic assays or the following assays, can be utilized to identify a subject having or at risk of developing a disorder associated with ORFX protein, nucleic acid expression or activity in, *e.g.*, proliferative or differentiative disorders such as hyperplasias, tumors, restenosis, psoriasis, Dupuytren's

contracture, diabetic complications, or rheumatoid arthritis, etc.; and glia-associated disorders such as cerebral lesions, diabetic neuropathies, cerebral edema, senile dementia, Alzheimer's disease, etc. Alternatively, the prognostic assays can be utilized to identify a subject having or at risk for developing a disease or disorder. Thus, the present invention provides a method for identifying a disease or disorder associated with aberrant ORFX expression or activity in which a test sample is obtained from a subject and ORFX protein or nucleic acid (*e.g.*, mRNA, genomic DNA) is detected, wherein the presence of ORFX protein or nucleic acid is diagnostic for a subject having or at risk of developing a disease or disorder associated with aberrant ORFX expression or activity. As used herein, a "test sample" refers to a biological sample obtained from a subject of interest. For example, a test sample can be a biological fluid (*e.g.*, serum), cell sample, or tissue.

Furthermore, the prognostic assays described herein can be used to determine whether a subject can be administered an agent (*e.g.*, an agonist, antagonist, peptidomimetic, protein, peptide, nucleic acid, small molecule, or other drug candidate) to treat a disease or disorder associated with aberrant ORFX expression or activity. For example, such methods can be used to determine whether a subject can be effectively treated with an agent for a disorder, such as a proliferative disorder, differentiative disorder, glia-associated disorders, etc. Thus, the present invention provides methods for determining whether a subject can be effectively treated with an agent for a disorder associated with aberrant ORFX expression or activity in which a test sample is obtained and ORFX protein or nucleic acid is detected (*e.g.*, wherein the presence of ORFX protein or nucleic acid is diagnostic for a subject that can be administered the agent to treat a disorder associated with aberrant ORFX expression or activity.)

The methods of the invention can also be used to detect genetic lesions in a ORFX gene, thereby determining if a subject with the lesioned gene is at risk for, or suffers from, a proliferative disorder, differentiative disorder, glia-associated disorder, etc. In various embodiments, the methods include detecting, in a sample of cells from the subject, the presence or absence of a genetic lesion characterized by at least one of an alteration affecting the integrity of a gene encoding a ORFX-protein, or the mis-expression of the ORFX gene. For example, such genetic lesions can be detected by ascertaining the existence of at least one of (1) a deletion of one or more nucleotides from a ORFX gene; (2) an addition of one or more nucleotides to a ORFX gene; (3) a substitution of one or more nucleotides of a ORFX gene, (4) a chromosomal

rearrangement of a ORFX gene; (5) an alteration in the level of a messenger RNA transcript of a ORFX gene, (6) aberrant modification of a ORFX gene, such as of the methylation pattern of the genomic DNA, (7) the presence of a non-wild type splicing pattern of a messenger RNA transcript of a ORFX gene, (8) a non-wild type level of a ORFX-protein, (9) allelic loss of a ORFX gene, and (10) inappropriate post-translational modification of a ORFX-protein. As described herein, there are a large number of assay techniques known in the art which can be used for detecting lesions in a ORFX gene. A preferred biological sample is a peripheral blood leukocyte sample isolated by conventional means from a subject. However, any biological sample containing nucleated cells may be used, including, for example, buccal mucosal cells.

In certain embodiments, detection of the lesion involves the use of a probe/primer in a polymerase chain reaction (PCR) (see, e.g., U.S. Pat. Nos. 4,683,195 and 4,683,202), such as anchor PCR or RACE PCR, or, alternatively, in a ligation chain reaction (LCR) (see, e.g., Landegran *et al.* (1988) *Science* 241:1077-1080; and Nakazawa *et al.* (1994) *PNAS* 91:360-364), the latter of which can be particularly useful for detecting point mutations in the ORFX-gene (see Abravaya *et al.* (1995) *Nucl Acids Res* 23:675-682). This method can include the steps of collecting a sample of cells from a patient, isolating nucleic acid (e.g., genomic, mRNA or both) from the cells of the sample, contacting the nucleic acid sample with one or more primers that specifically hybridize to a ORFX gene under conditions such that hybridization and amplification of the ORFX gene (if present) occurs, and detecting the presence or absence of an amplification product, or detecting the size of the amplification product and comparing the length to a control sample. It is anticipated that PCR and/or LCR may be desirable to use as a preliminary amplification step in conjunction with any of the techniques used for detecting mutations described herein.

Alternative amplification methods include: self sustained sequence replication (Guatelli *et al.*, 1990, *Proc Natl Acad Sci USA* 87:1874-1878), transcriptional amplification system (Kwoh, *et al.*, 1989, *Proc Natl Acad Sci USA* 86:1173-1177), Q-Beta Replicase (Lizardi *et al.*, 1988, *BioTechnology* 6:1197), or any other nucleic acid amplification method, followed by the detection of the amplified molecules using techniques well known to those of skill in the art. These detection schemes are especially useful for the detection of nucleic acid molecules if such molecules are present in very low numbers.

In an alternative embodiment, mutations in a ORFX gene from a sample cell can be identified by alterations in restriction enzyme cleavage patterns. For example, sample and control DNA is isolated, amplified (optionally), digested with one or more restriction endonucleases, and fragment length sizes are determined by gel electrophoresis and compared.

5 Differences in fragment length sizes between sample and control DNA indicates mutations in the sample DNA. Moreover, the use of sequence specific ribozymes (see, for example, U.S. Pat. No. 5,493,531) can be used to score for the presence of specific mutations by development or loss of a ribozyme cleavage site.

In other embodiments, genetic mutations in ORFX can be identified by hybridizing a

10 sample and control nucleic acids, e.g., DNA or RNA, to high density arrays containing hundreds or thousands of oligonucleotide probes (Cronin *et al.* (1996) *Human Mutation* 7: 244-255; Kozal *et al.* (1996) *Nature Medicine* 2: 753-759). For example, genetic mutations in ORFX can be identified in two dimensional arrays containing light-generated DNA probes as described in Cronin *et al.* above. Briefly, a first hybridization array of probes can be used to scan through

15 long stretches of DNA in a sample and control to identify base changes between the sequences by making linear arrays of sequential overlapping probes. This step allows the identification of point mutations. This step is followed by a second hybridization array that allows the characterization of specific mutations by using smaller, specialized probe arrays complementary to all variants or mutations detected. Each mutation array is composed of parallel probe sets, one

20 complementary to the wild-type gene and the other complementary to the mutant gene.

In yet another embodiment, any of a variety of sequencing reactions known in the art can be used to directly sequence the ORFX gene and detect mutations by comparing the sequence of the sample ORFX with the corresponding wild-type (control) sequence. Examples of sequencing reactions include those based on techniques developed by Maxim and Gilbert (1977) *PNAS*

25 74:560 or Sanger (1977) *PNAS* 74:5463. It is also contemplated that any of a variety of automated sequencing procedures can be utilized when performing the diagnostic assays (Naeve *et al.*, (1995) *Biotechniques* 19:448), including sequencing by mass spectrometry (see, e.g., PCT International Publ. No. WO 94/16101; Cohen *et al.* (1996) *Adv Chromatogr* 36:127-162; and Griffin *et al.* (1993) *Appl Biochem Biotechnol* 38:147-159).

30 Other methods for detecting mutations in the ORFX gene include methods in which protection from cleavage agents is used to detect mismatched bases in RNA/RNA or RNA/DNA

heteroduplexes (Myers *et al.* (1985) *Science* 230:1242). In general, the art technique of "mismatch cleavage" starts by providing heteroduplexes of formed by hybridizing (labeled) RNA or DNA containing the wild-type ORFX sequence with potentially mutant RNA or DNA obtained from a tissue sample. The double-stranded duplexes are treated with an agent that cleaves single-stranded regions of the duplex such as which will exist due to basepair mismatches between the control and sample strands. For instance, RNA/DNA duplexes can be treated with RNase and DNA/DNA hybrids treated with S1 nuclease to enzymatically digest the mismatched regions. In other embodiments, either DNA/DNA or RNA/DNA duplexes can be treated with hydroxylamine or osmium tetroxide and with piperidine in order to digest mismatched regions. After digestion of the mismatched regions, the resulting material is then separated by size on denaturing polyacrylamide gels to determine the site of mutation. See, for example, Cotton *et al* (1988) *Proc Natl Acad Sci USA* 85:4397; Saleeba *et al* (1992) *Methods Enzymol* 217:286-295. In an embodiment, the control DNA or RNA can be labeled for detection.

In still another embodiment, the mismatch cleavage reaction employs one or more proteins that recognize mismatched base pairs in double-stranded DNA (so called "DNA mismatch repair" enzymes) in defined systems for detecting and mapping point mutations in ORFX cDNAs obtained from samples of cells. For example, the mutY enzyme of *E. coli* cleaves A at G/A mismatches and the thymidine DNA glycosylase from HeLa cells cleaves T at G/T mismatches (Hsu *et al.* (1994) *Carcinogenesis* 15:1657-1662). According to an exemplary embodiment, a probe based on a ORFX sequence, *e.g.*, a wild-type ORFX sequence, is hybridized to a cDNA or other DNA product from a test cell(s). The duplex is treated with a DNA mismatch repair enzyme, and the cleavage products, if any, can be detected from electrophoresis protocols or the like. See, for example, U.S. Pat. No. 5,459,039.

In other embodiments, alterations in electrophoretic mobility will be used to identify mutations in ORFX genes. For example, single strand conformation polymorphism (SSCP) may be used to detect differences in electrophoretic mobility between mutant and wild type nucleic acids (Orita *et al.* (1989) *Proc Natl Acad Sci USA*: 86:2766, see also Cotton (1993) *Mutat Res* 285:125-144; Hayashi (1992) *Genet Anal Tech Appl* 9:73-79). Single-stranded DNA fragments of sample and control ORFX nucleic acids will be denatured and allowed to renature. The secondary structure of single-stranded nucleic acids varies according to sequence, the resulting

alteration in electrophoretic mobility enables the detection of even a single base change. The DNA fragments may be labeled or detected with labeled probes. The sensitivity of the assay may be enhanced by using RNA, rather than DNA, in which the secondary structure is more sensitive to a change in sequence. In one embodiment, the subject method utilizes heteroduplex analysis to separate double stranded heteroduplex molecules on the basis of changes in electrophoretic mobility. See, e.g., Keen *et al.* (1991) *Trends Genet* 7:5.

In yet another embodiment the movement of mutant or wild-type fragments in polyacrylamide gels containing a gradient of denaturant is assayed using denaturing gradient gel electrophoresis (DGGE). See, e.g., Myers *et al.* (1985) *Nature* 313:495. When DGGE is used as the method of analysis, DNA will be modified to insure that it does not completely denature, for example by adding a GC clamp of approximately 40 bp of high-melting GC-rich DNA by PCR. In a further embodiment, a temperature gradient is used in place of a denaturing gradient to identify differences in the mobility of control and sample DNA. See, e.g., Rosenbaum and Reissner (1987) *Biophys Chem* 265:12753.

Examples of other techniques for detecting point mutations include, but are not limited to, selective oligonucleotide hybridization, selective amplification, or selective primer extension. For example, oligonucleotide primers may be prepared in which the known mutation is placed centrally and then hybridized to target DNA under conditions that permit hybridization only if a perfect match is found. See, e.g., Saiki *et al.* (1986) *Nature* 324:163; Saiki *et al.* (1989) *Proc Natl Acad. Sci USA* 86:6230. Such allele specific oligonucleotides are hybridized to PCR amplified target DNA or a number of different mutations when the oligonucleotides are attached to the hybridizing membrane and hybridized with labeled target DNA.

Alternatively, allele specific amplification technology that depends on selective PCR amplification may be used in conjunction with the instant invention. Oligonucleotides used as primers for specific amplification may carry the mutation of interest in the center of the molecule (so that amplification depends on differential hybridization) (Gibbs *et al.* (1989) *Nucleic Acids Res* 17:2437-2448) or at the extreme 3' end of one primer where, under appropriate conditions, mismatch can prevent, or reduce polymerase extension (Prossner (1993) *Tibtech* 11:238). In addition it may be desirable to introduce a novel restriction site in the region of the mutation to create cleavage-based detection. See, e.g., Gasparini *et al.* (1992) *Mol Cell Probes* 6:1. It is anticipated that in certain embodiments amplification may also be performed using Taq ligase

for amplification. See, e.g., Barany (1991) *Proc Natl Acad Sci USA* 88:189. In such cases, ligation will occur only if there is a perfect match at the 3' end of the 5' sequence, making it possible to detect the presence of a known mutation at a specific site by looking for the presence or absence of amplification.

The methods described herein may be performed, for example, by utilizing pre-packaged diagnostic kits comprising at least one probe nucleic acid or antibody reagent described herein, which may be conveniently used, e.g., in clinical settings to diagnose patients exhibiting symptoms or family history of a disease or illness involving a ORFX gene.

Furthermore, any cell type or tissue, preferably peripheral blood leukocytes, in which ORFX is expressed may be utilized in the prognostic assays described herein. However, any biological sample containing nucleated cells may be used, including, for example, buccal mucosal cells.

Pharmacogenomics

Agents, or modulators that have a stimulatory or inhibitory effect on ORFX activity (e.g., ORFX gene expression), as identified by a screening assay described herein can be administered to individuals to treat (prophylactically or therapeutically) disorders (e.g., neurological, cancer-related or gestational disorders) associated with aberrant ORFX activity. In conjunction with such treatment, the pharmacogenomics (i.e., the study of the relationship between an individual's genotype and that individual's response to a foreign compound or drug) of the individual may be considered. Differences in metabolism of therapeutics can lead to severe toxicity or therapeutic failure by altering the relation between dose and blood concentration of the pharmacologically active drug. Thus, the pharmacogenomics of the individual permits the selection of effective agents (e.g., drugs) for prophylactic or therapeutic treatments based on a consideration of the individual's genotype. Such pharmacogenomics can further be used to determine appropriate dosages and therapeutic regimens. Accordingly, the activity of ORFX protein, expression of ORFX nucleic acid, or mutation content of ORFX genes in an individual can be determined to thereby select appropriate agent(s) for therapeutic or prophylactic treatment of the individual.

Pharmacogenomics deals with clinically significant hereditary variations in the response to drugs due to altered drug disposition and abnormal action in affected persons. See e.g., Eichelbaum, 1996, *Clin Exp Pharmacol Physiol*, 23:983-985 and Linder, 1997, *Clin Chem*, 43:254-266. In general, two types of pharmacogenetic conditions can be differentiated. Genetic

conditions transmitted as a single factor altering the way drugs act on the body (altered drug action) or genetic conditions transmitted as single factors altering the way the body acts on drugs (altered drug metabolism). These pharmacogenetic conditions can occur either as rare defects or as polymorphisms. For example, glucose-6-phosphate dehydrogenase (G6PD) deficiency is a common inherited enzymopathy in which the main clinical complication is haemolysis after ingestion of oxidant drugs (anti-malarials, sulfonamides, analgesics, nitrofurans) and consumption of fava beans.

As an illustrative embodiment, the activity of drug metabolizing enzymes is a major determinant of both the intensity and duration of drug action. The discovery of genetic polymorphisms of drug metabolizing enzymes (*e.g.*, N-acetyltransferase 2 (NAT 2) and cytochrome P450 enzymes CYP2D6 and CYP2C19) has provided an explanation as to why some patients do not obtain the expected drug effects or show exaggerated drug response and serious toxicity after taking the standard and safe dose of a drug. These polymorphisms are expressed in two phenotypes in the population, the extensive metabolizer (EM) and poor metabolizer (PM). The prevalence of PM is different among different populations. For example, the gene coding for CYP2D6 is highly polymorphic and several mutations have been identified in PM, which all lead to the absence of functional CYP2D6. Poor metabolizers of CYP2D6 and CYP2C19 quite frequently experience exaggerated drug response and side effects when they receive standard doses. If a metabolite is the active therapeutic moiety, PM show no therapeutic response, as demonstrated for the analgesic effect of codeine mediated by its CYP2D6-formed metabolite morphine. The other extreme are the so called ultra-rapid metabolizers who do not respond to standard doses. Recently, the molecular basis of ultra-rapid metabolism has been identified to be due to CYP2D6 gene amplification.

Thus, the activity of ORFX protein, expression of ORFX nucleic acid, or mutation content of ORFX genes in an individual can be determined to thereby select appropriate agent(s) for therapeutic or prophylactic treatment of the individual. In addition, pharmacogenetic studies can be used to apply genotyping of polymorphic alleles encoding drug-metabolizing enzymes to the identification of an individual's drug responsiveness phenotype. This knowledge, when applied to dosing or drug selection, can avoid adverse reactions or therapeutic failure and thus enhance therapeutic or prophylactic efficiency when treating a subject with a ORFX modulator, such as a modulator identified by one of the exemplary screening assays described herein.

Monitoring Clinical Efficacy

Monitoring the influence of agents (*e.g.*, drugs, compounds) on the expression or activity of ORFX (*e.g.*, the ability to modulate aberrant cell proliferation and/or differentiation) can be applied in basic drug screening and in clinical trials. For example, the effectiveness of an agent determined by a screening assay as described herein to increase ORFX gene expression, protein levels, or upregulate ORFX activity, can be monitored in clinical trials of subjects exhibiting decreased ORFX gene expression, protein levels, or downregulated ORFX activity.

Alternatively, the effectiveness of an agent determined by a screening assay to decrease ORFX gene expression, protein levels, or downregulate ORFX activity, can be monitored in clinical trials of subjects exhibiting increased ORFX gene expression, protein levels, or upregulated ORFX activity. In such clinical trials, the expression or activity of ORFX and, preferably, other genes that have been implicated in, for example, a proliferative or neurological disorder, can be used as a "read out" or marker of the responsiveness of a particular cell.

For example, genes, including ORFX, that are modulated in cells by treatment with an agent (*e.g.*, compound, drug or small molecule) that modulates ORFX activity (*e.g.*, identified in a screening assay as described herein) can be identified. Thus, to study the effect of agents on cellular proliferation disorders, for example, in a clinical trial, cells can be isolated and RNA prepared and analyzed for the levels of expression of ORFX and other genes implicated in the disorder. The levels of gene expression (*i.e.*, a gene expression pattern) can be quantified by Northern blot analysis or RT-PCR, as described herein, or alternatively by measuring the amount of protein produced, by one of the methods as described herein, or by measuring the levels of activity of ORFX or other genes. In this way, the gene expression pattern can serve as a marker, indicative of the physiological response of the cells to the agent. Accordingly, this response state may be determined before, and at various points during, treatment of the individual with the agent.

In one embodiment, the invention provides a method for monitoring the effectiveness of treatment of a subject with an agent (*e.g.*, an agonist, antagonist, protein, peptide, nucleic acid, peptidomimetic, small molecule, or other drug candidate identified by the screening assays described herein) comprising the steps of (i) obtaining a pre-administration sample from a subject prior to administration of the agent; (ii) detecting the level of expression of a ORFX protein, mRNA, or genomic DNA in the preadministration sample; (iii) obtaining one or more

post-administration samples from the subject; (iv) detecting the level of expression or activity of the ORFX protein, mRNA, or genomic DNA in the post-administration samples; (v) comparing the level of expression or activity of the ORFX protein, mRNA, or genomic DNA in the pre-administration sample with the ORFX protein, mRNA, or genomic DNA in the post administration sample or samples; and (vi) altering the administration of the agent to the subject accordingly. For example, increased administration of the agent may be desirable to increase the expression or activity of ORFX to higher levels than detected, *i.e.*, to increase the effectiveness of the agent. Alternatively, decreased administration of the agent may be desirable to decrease expression or activity of ORFX to lower levels than detected, *i.e.*, to decrease the effectiveness of the agent.

Methods of Treatment

The present invention provides for both prophylactic and therapeutic methods of treating a subject at risk of (or susceptible to) a disorder or having a disorder associated with aberrant ORFX expression or activity.

Diseases and disorders that are characterized by increased (relative to a subject not suffering from the disease or disorder) levels or biological activity may be treated with Therapeutics that antagonize (*i.e.*, reduce or inhibit) activity. Therapeutics that antagonize activity may be administered in a therapeutic or prophylactic manner. Therapeutics that may be utilized include, but are not limited to, (i) a ORFX polypeptide, or analogs, derivatives, fragments or homologs thereof; (ii) antibodies to a ORFX peptide; (iii) nucleic acids encoding a ORFX peptide; (iv) administration of antisense nucleic acid and nucleic acids that are "dysfunctional" (*i.e.*, due to a heterologous insertion within the coding sequences of coding sequences to a ORFX peptide) that are utilized to "knockout" endogenous function of a ORFX peptide by homologous recombination (see, *e.g.*, Capecchi, 1989, *Science* 244: 1288-1292); or (v) modulators (*i.e.*, inhibitors, agonists and antagonists, including additional peptide mimetic of the invention or antibodies specific to a peptide of the invention) that alter the interaction between a ORFX peptide and its binding partner.

Diseases and disorders that are characterized by decreased (relative to a subject not suffering from the disease or disorder) levels or biological activity may be treated with Therapeutics that increase (*i.e.*, are agonists to) activity. Therapeutics that upregulate activity may be administered in a therapeutic or prophylactic manner. Therapeutics that may be utilized

include, but are not limited to, a ORFX peptide, or analogs, derivatives, fragments or homologs thereof; or an agonist that increases bioavailability.

Increased or decreased levels can be readily detected by quantifying peptide and/or RNA, by obtaining a patient tissue sample (e.g., from biopsy tissue) and assaying it *in vitro* for RNA or peptide levels, structure and/or activity of the expressed peptides (or mRNAs of a ORFX peptide). Methods that are well-known within the art include, but are not limited to, immunoassays (e.g., by Western blot analysis, immunoprecipitation followed by sodium dodecyl sulfate (SDS) polyacrylamide gel electrophoresis, immunocytochemistry, etc.) and/or hybridization assays to detect expression of mRNAs (e.g., Northern assays, dot blots, *in situ* hybridization, etc.).

In one aspect, the invention provides a method for preventing, in a subject, a disease or condition associated with an aberrant ORFX expression or activity, by administering to the subject an agent that modulates ORFX expression or at least one ORFX activity. Subjects at risk for a disease that is caused or contributed to by aberrant ORFX expression or activity can be identified by, for example, any or a combination of diagnostic or prognostic assays as described herein. Administration of a prophylactic agent can occur prior to the manifestation of symptoms characteristic of the ORFX aberrancy, such that a disease or disorder is prevented or, alternatively, delayed in its progression. Depending on the type of ORFX aberrancy, for example, a ORFX agonist or ORFX antagonist agent can be used for treating the subject. The appropriate agent can be determined based on screening assays described herein.

Another aspect of the invention pertains to methods of modulating ORFX expression or activity for therapeutic purposes. The modulatory method of the invention involves contacting a cell with an agent that modulates one or more of the activities of ORFX protein activity associated with the cell. An agent that modulates ORFX protein activity can be an agent as described herein, such as a nucleic acid or a protein, a naturally-occurring cognate ligand of a ORFX protein, a peptide, a ORFX peptidomimetic, or other small molecule. In one embodiment, the agent stimulates one or more ORFX protein activity. Examples of such stimulatory agents include active ORFX protein and a nucleic acid molecule encoding ORFX that has been introduced into the cell. In another embodiment, the agent inhibits one or more ORFX protein activity. Examples of such inhibitory agents include antisense ORFX nucleic acid molecules and anti-ORFX antibodies. These modulatory methods can be performed *in vitro*

(*e.g.*, by culturing the cell with the agent) or, alternatively, *in vivo* (*e.g.*, by administering the agent to a subject). As such, the present invention provides methods of treating an individual afflicted with a disease or disorder characterized by aberrant expression or activity of a ORFX protein or nucleic acid molecule. In one embodiment, the method involves administering an agent (*e.g.*, an agent identified by a screening assay described herein), or combination of agents that modulates (*e.g.*, upregulates or downregulates) ORFX expression or activity. In another embodiment, the method involves administering a ORFX protein or nucleic acid molecule as therapy to compensate for reduced or aberrant ORFX expression or activity.

Determination of the Biological Effect of a Therapeutic

In various embodiments of the present invention, suitable *in vitro* or *in vivo* assays are utilized to determine the effect of a specific Therapeutic and whether its administration is indicated for treatment of the affected tissue.

In various specific embodiments, *in vitro* assays may be performed with representative cells of the type(s) involved in the patient's disorder, to determine if a given Therapeutic exerts the desired effect upon the cell type(s). Compounds for use in therapy may be tested in suitable animal model systems including, but not limited to rats, mice, chicken, cows, monkeys, rabbits, and the like, prior to testing in human subjects. Similarly, for *in vivo* testing, any of the animal model system known in the art may be used prior to administration to human subjects.

Malignancies

Some ORFX polypeptides are expressed in cancerous cells (*see, e.g.*, Tables 1 and 2). Accordingly, the corresponding ORF protein is involved in the regulation of cell proliferation. Accordingly, Therapeutics of the present invention may be useful in the therapeutic or prophylactic treatment of diseases or disorders that are associated with cell hyperproliferation and/or loss of control of cell proliferation (*e.g.*, cancers, malignancies and tumors). For a review of such hyperproliferation disorders, *see e.g.*, Fishman, *et al.*, 1985. MEDICINE, 2nd ed., J.B. Lippincott Co., Philadelphia, PA.

Therapeutics of the present invention may be assayed by any method known within the art for efficacy in treating or preventing malignancies and related disorders. Such assays include,

but are not limited to, *in vitro* assays utilizing transformed cells or cells derived from the patient's tumor, as well as *in vivo* assays using animal models of cancer or malignancies. Potentially effective Therapeutics are those that, for example, inhibit the proliferation of tumor-derived or transformed cells in culture or cause a regression of tumors in animal models, in comparison to the controls.

In the practice of the present invention, once a malignancy or cancer has been shown to be amenable to treatment by modulating (*i.e.*, inhibiting, antagonizing or agonizing) activity, that cancer or malignancy may subsequently be treated or prevented by the administration of a Therapeutic that serves to modulate protein function.

Premalignant conditions

The Therapeutics of the present invention that are effective in the therapeutic or prophylactic treatment of cancer or malignancies may also be administered for the treatment of pre-malignant conditions and/or to prevent the progression of a pre-malignancy to a neoplastic or malignant state. Such prophylactic or therapeutic use is indicated in conditions known or suspected of preceding progression to neoplasia or cancer, in particular, where non-neoplastic cell growth consisting of hyperplasia, metaplasia or, most particularly, dysplasia has occurred. For a review of such abnormal cell growth see *e.g.*, Robbins & Angell, 1976. BASIC PATHOLOGY, 2nd ed., W.B. Saunders Co., Philadelphia, PA.

Hyperplasia is a form of controlled cell proliferation involving an increase in cell number in a tissue or organ, without significant alteration in its structure or function. For example, it has been demonstrated that endometrial hyperplasia often precedes endometrial cancer. Metaplasia is a form of controlled cell growth in which one type of mature or fully differentiated cell substitutes for another type of mature cell. Metaplasia may occur in epithelial or connective tissue cells. Dysplasia is generally considered a precursor of cancer, and is found mainly in the epithelia. Dysplasia is the most disorderly form of non-neoplastic cell growth, and involves a loss in individual cell uniformity and in the architectural orientation of cells. Dysplasia characteristically occurs where there exists chronic irritation or inflammation, and is often found in the cervix, respiratory passages, oral cavity, and gall bladder.

Alternatively, or in addition to the presence of abnormal cell growth characterized as hyperplasia, metaplasia, or dysplasia, the presence of one or more characteristics of a transformed or malignant phenotype displayed either *in vivo* or *in vitro* within a cell sample derived from a patient, is indicative of the desirability of prophylactic/therapeutic administration of a Therapeutic that possesses the ability to modulate activity of An aforementioned protein. Characteristics of a transformed phenotype include, but are not limited to: (i) morphological changes; (ii) looser substratum attachment; (iii) loss of cell-to-cell contact inhibition; (iv) loss of anchorage dependence; (v) protease release; (vi) increased sugar transport; (vii) decreased serum requirement; (viii) expression of fetal antigens, (ix) disappearance of the 250 kDal cell-surface protein, and the like. See *e.g.*, Richards, *et al.*, 1986. MOLECULAR PATHOLOGY, W.B. Saunders Co., Philadelphia, PA.

In a specific embodiment of the present invention, a patient that exhibits one or more of the following predisposing factors for malignancy is treated by administration of an effective amount of a Therapeutic: (i) a chromosomal translocation associated with a malignancy (*e.g.*, the Philadelphia chromosome (*bcr/abl*) for chronic myelogenous leukemia and t(14;18) for follicular lymphoma, etc.); (ii) familial polyposis or Gardner's syndrome (possible forerunners of colon cancer); (iii) monoclonal gammopathy of undetermined significance (a possible precursor of multiple myeloma) and (iv) a first degree kinship with persons having a cancer or pre-cancerous disease showing a Mendelian (genetic) inheritance pattern (*e.g.*, familial polyposis of the colon, Gardner's syndrome, hereditary exostosis, polyendocrine adenomatosis, Peutz-Jeghers syndrome, neurofibromatosis of Von Recklinghausen, medullary thyroid carcinoma with amyloid production and pheochromocytoma, retinoblastoma, carotid body tumor, cutaneous melanocarcinoma, intraocular melanocarcinoma, xeroderma pigmentosum, ataxia telangiectasia, Chediak-Higashi syndrome, albinism, Fanconi's aplastic anemia and Bloom's syndrome).

In another embodiment, a Therapeutic of the present invention is administered to a human patient to prevent the progression to breast, colon, lung, pancreatic, or uterine cancer, or melanoma or sarcoma.

Hyperproliferative and dysproliferative disorders

In one embodiment of the present invention, a Therapeutic is administered in the therapeutic or prophylactic treatment of hyperproliferative or benign dysproliferative disorders. The efficacy in treating or preventing hyperproliferative diseases or disorders of a Therapeutic of the present invention may be assayed by any method known within the art. Such assays include
5 *in vitro* cell proliferation assays, *in vitro* or *in vivo* assays using animal models of hyperproliferative diseases or disorders, or the like. Potentially effective Therapeutics may, for example, promote cell proliferation in culture or cause growth or cell proliferation in animal models in comparison to controls.

Specific embodiments of the present invention are directed to the treatment or prevention of cirrhosis of the liver (a condition in which scarring has overtaken normal liver regeneration processes); treatment of keloid (hypertrophic scar) formation causing disfiguring of the skin in which the scarring process interferes with normal renewal; psoriasis (a common skin condition characterized by excessive proliferation of the skin and delay in proper cell fate determination);
15 benign tumors; fibrocystic conditions and tissue hypertrophy (*e.g.*, benign prostatic hypertrophy).

Neurodegenerative disorders

Some ORFX proteins are found in cell types have been implicated in the deregulation of cellular maturation and apoptosis, which are both characteristic of neurodegenerative disease. Accordingly, Therapeutics of the invention, particularly but not limited to those that modulate
20 (or supply) activity of an aforementioned protein, may be effective in treating or preventing neurodegenerative disease. Therapeutics of the present invention that modulate the activity of an aforementioned protein involved in neurodegenerative disorders can be assayed by any method known in the art for efficacy in treating or preventing such neurodegenerative diseases and disorders. Such assays include *in vitro* assays for regulated cell maturation or inhibition of
25 apoptosis or *in vivo* assays using animal models of neurodegenerative diseases or disorders, or any of the assays described below. Potentially effective Therapeutics, for example but not by way of limitation, promote regulated cell maturation and prevent cell apoptosis in culture, or reduce neurodegeneration in animal models in comparison to controls.

Once a neurodegenerative disease or disorder has been shown to be amenable to treatment by modulation activity, that neurodegenerative disease or disorder can be treated or prevented by administration of a Therapeutic that modulates activity. Such diseases include all degenerative disorders involved with aging, especially osteoarthritis and neurodegenerative disorders.

Disorders related to organ transplantation

Some ORFX can be associated with disorders related to organ transplantation, in particular but not limited to organ rejection. Therapeutics of the invention, particularly those that modulate (or supply) activity, may be effective in treating or preventing diseases or disorders related to organ transplantation. Therapeutics of the invention (particularly Therapeutics that modulate the levels or activity of an aforementioned protein) can be assayed by any method known in the art for efficacy in treating or preventing such diseases and disorders related to organ transplantation. Such assays include *in vitro* assays for using cell culture models as described below, or *in vivo* assays using animal models of diseases and disorders related to organ transplantation, see *e.g.*, below. Potentially effective Therapeutics, for example but not by way of limitation, reduce immune rejection responses in animal models in comparison to controls.

Accordingly, once diseases and disorders related to organ transplantation are shown to be amenable to treatment by modulation of activity, such diseases or disorders can be treated or prevented by administration of a Therapeutic that modulates activity.

Cardiovascular Disease

GENX has been implicated in cardiovascular disorders, including in atherosclerotic plaque formation. Diseases such as cardiovascular disease, including cerebral thrombosis or hemorrhage, ischemic heart or renal disease, peripheral vascular disease, or thrombosis of other major vessel, and other diseases, including diabetes mellitus, hypertension, hypothyroidism, cholesterol ester storage disease, systemic lupus erythematosus, homocysteinemia, and familial protein or lipid processing diseases, and the like, are either directly or indirectly associated with atherosclerosis. Accordingly, Therapeutics of the invention, particularly those that modulate (or supply) activity or formation may be effective in treating or preventing

atherosclerosis-associated diseases or disorders. Therapeutics of the invention (particularly Therapeutics that modulate the levels or activity) can be assayed by any method known in the art, including those described below, for efficacy in treating or preventing such diseases and disorders.

5 A vast array of animal and cell culture models exist for processes involved in atherosclerosis. A limited and non-exclusive list of animal models includes knockout mice for premature atherosclerosis (Kurabayashi and Yazaki, 1996, *Int. Angiol.* 15: 187-194), transgenic mouse models of atherosclerosis (Kappel *et al.*, 1994, *FASEB J.* 8: 583-592), antisense oligonucleotide treatment of animal models (Callow, 1995, *Curr. Opin. Cardiol.* 10: 569-576),
10 transgenic rabbit models for atherosclerosis (Taylor, 1997, *Ann. N.Y. Acad. Sci.* 811: 146-152), hypercholesterolemic animal models (Rosenfeld, 1996, *Diabetes Res. Clin. Pract.* 30 Suppl.: 1-11), hyperlipidemic mice (Paigen *et al.*, 1994, *Curr. Opin. Lipidol.* 5: 258-264), and inhibition of lipoxygenase in animals (Sigal *et al.*, 1994, *Ann. N.Y. Acad. Sci.* 714: 211-224). In addition, *in vitro* cell models include but are not limited to monocytes exposed to low density lipoprotein
15 (Frostegard *et al.*, 1996, *Atherosclerosis* 121: 93-103), cloned vascular smooth muscle cells (Suttles *et al.*, 1995, *Exp. Cell Res.* 218: 331-338), endothelial cell-derived chemoattractant exposed T cells (Katz *et al.*, 1994, *J. Leukoc. Biol.* 55: 567-573), cultured human aortic endothelial cells (Farber *et al.*, 1992, *Am. J. Physiol.* 262: H1088-1085), and foam cell cultures (Libby *et al.*, 1996, *Curr Opin Lipidol* 7: 330-335). Potentially effective Therapeutics, for
20 example but not by way of limitation, reduce foam cell formation in cell culture models, or reduce atherosclerotic plaque formation in hypercholesterolemic mouse models of atherosclerosis in comparison to controls.

Accordingly, once an atherosclerosis-associated disease or disorder has been shown to be amenable to treatment by modulation of activity or formation, that disease or disorder can be
25 treated or prevented by administration of a Therapeutic that modulates activity.

Cytokine and Cell Proliferation/Differentiation Activity

A GENX protein of the present invention may exhibit cytokine, cell proliferation (either inducing or inhibiting) or cell differentiation (either inducing or inhibiting) activity or may induce production of other cytokines in certain cell populations. Many protein factors discovered

to date, including all known cytokines, have exhibited activity in one or more factor dependent cell proliferation assays, and hence the assays serve as a convenient confirmation of cytokine activity. The activity of a protein of the present invention is evidenced by any one of a number of routine factor dependent cell proliferation assays for cell lines including, without limitation, 32D,
5 DA2, DA1G, T10, B9, B9/11, BaF3, MC9/G, M+ (preB M+), 2E8, RB5, DA1, 123, T1165, HT2, CTLL2, TF-1, Mo7e and CMK.

The activity of a protein of the invention may, among other means, be measured by the following methods: Assays for T-cell or thymocyte proliferation include without limitation those described in: CURRENT PROTOCOLS IN IMMUNOLOGY, Ed by Coligan *et al.*, Greene Publishing
10 Associates and Wiley-Interscience (Chapter 3 and Chapter 7); Takai *et al.*, *J Immunol* 137:3494-3500, 1986; Bertagnoli *et al.*, *J Immunol* 145:1706-1712, 1990; Bertagnoli *et al.*, *Cell Immunol* 133:327-341, 1991; Bertagnoli, *et al.*, *J Immunol* 149:3778-3783, 1992; Bowman *et al.*, *J Immunol* 152:1756-1761, 1994.

Assays for cytokine production and/or proliferation of spleen cells, lymph node cells or
15 thymocytes include, without limitation, those described by Kruisbeek and Shevach, In: CURRENT PROTOCOLS IN IMMUNOLOGY. Coligan *et al.*, eds. Vol 1, pp. 3.12.1-14, John Wiley and Sons, Toronto 1994; and by Schreiber, In: CURRENT PROTOCOLS IN IMMUNOLOGY. Coligan eds. Vol 1 pp. 6.8.1-8, John Wiley and Sons, Toronto 1994.

Assays for proliferation and differentiation of hematopoietic and lymphopoietic cells
20 include, without limitation, those described by Bottomly *et al.*, In: CURRENT PROTOCOLS IN IMMUNOLOGY. Coligan *et al.*, eds. Vol 1 pp. 6.3.1-6.3.12, John Wiley and Sons, Toronto 1991; deVries *et al.*, *J Exp Med* 173:1205-1211, 1991; Moreau *et al.*, *Nature* 336:690-692, 1988; Greenberger *et al.*, *Proc Natl Acad Sci U.S.A.* 80:2931-2938, 1983; Nordan, In: CURRENT PROTOCOLS IN IMMUNOLOGY. Coligan *et al.*, eds. Vol 1 pp. 6.6.1-5, John Wiley and Sons,
25 Toronto 1991; Smith *et al.*, *Proc Natl Acad Sci U.S.A.* 83:1857-1861, 1986; Measurement of human Interleukin 11-Bennett, *et al.* In: CURRENT PROTOCOLS IN IMMUNOLOGY. Coligan *et al.*, eds. Vol 1 pp. 6.15.1 John Wiley and Sons, Toronto 1991; Ciarletta, *et al.*, In: CURRENT PROTOCOLS IN IMMUNOLOGY. Coligan *et al.*, eds. Vol 1 pp. 6.13.1, John Wiley and Sons, Toronto 1991.

30 Assays for T-cell clone responses to antigens (which will identify, among others, proteins that affect APC-T cell interactions as well as direct T-cell effects by measuring proliferation and

cytokine production) include, without limitation, those described in: CURRENT PROTOCOLS IN IMMUNOLOGY. Coligan *et al.*, eds., Greene Publishing Associates and Wiley-Interscience (Chapter 3, Chapter 6, Chapter 7); Weinberger *et al.*, *Proc Natl Acad Sci USA* 77:6091-6095, 1980; Weinberger *et al.*, *Eur J Immunol* 11:405-411, 1981; Takai *et al.*, *J Immunol* 137:3494-3500, 1986; Takai *et al.*, *J Immunol* 140:508-512, 1988.

Immune Stimulating or Suppressing Activity

A GENX protein of the present invention may also exhibit immune stimulating or immune suppressing activity, including without limitation the activities for which assays are described herein. A protein may be useful in the treatment of various immune deficiencies and disorders (including severe combined immunodeficiency (SCID)), *e.g.*, in regulating (up or down) growth and proliferation of T and/or B lymphocytes, as well as effecting the cytolytic activity of NK cells and other cell populations. These immune deficiencies may be genetic or be caused by viral (*e.g.*, HIV) as well as bacterial or fungal infections, or may result from autoimmune disorders. More specifically, infectious diseases caused by viral, bacterial, fungal or other infection may be treatable using a protein of the present invention, including infections by HIV, hepatitis viruses, herpesviruses, mycobacteria, Leishmania species., malaria species. and various fungal infections such as candidiasis. Of course, in this regard, a protein of the present invention may also be useful where a boost to the immune system generally may be desirable, *i.e.*, in the treatment of cancer.

Autoimmune disorders which may be treated using a protein of the present invention include, for example, connective tissue disease, multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation, Guillain-Barre syndrome, autoimmune thyroiditis, insulin dependent diabetes mellitus, myasthenia gravis, graft-versus-host disease and autoimmune inflammatory eye disease. Such a protein of the present invention may also be useful in the treatment of allergic reactions and conditions, such as asthma (particularly allergic asthma) or other respiratory problems. Other conditions, in which immune suppression is desired (including, for example, organ transplantation), may also be treatable using a protein of the present invention.

Using the proteins of the invention it may also be possible to immune responses, in a number of ways. Down regulation may be in the form of inhibiting or blocking an immune response already in progress or may involve preventing the induction of an immune response.

The functions of activated T cells may be inhibited by suppressing T cell responses or by inducing specific tolerance in T cells, or both. Immunosuppression of T cell responses is generally an active, non-antigen-specific, process which requires continuous exposure of the T cells to the suppressive agent. Tolerance, which involves inducing non-responsiveness or energy
5 in T cells, is distinguishable from immunosuppression in that it is generally antigen-specific and persists after exposure to the tolerizing agent has ceased. Operationally, tolerance can be demonstrated by the lack of a T cell response upon re-exposure to specific antigen in the absence of the tolerizing agent.

Down regulating or preventing one or more antigen functions (including without
10 limitation B lymphocyte antigen functions (such as, for example, B7), *e.g.*, preventing high level lymphokine synthesis by activated T cells, will be useful in situations of tissue, skin and organ transplantation and in graft-versus-host disease (GVHD). For example, blockage of T cell function should result in reduced tissue destruction in tissue transplantation. Typically, in tissue transplants, rejection of the transplant is initiated through its recognition as foreign by T cells,
15 followed by an immune reaction that destroys the transplant. The administration of a molecule which inhibits or blocks interaction of a B7 lymphocyte antigen with its natural ligand(s) on immune cells (such as a soluble, monomeric form of a peptide having B7-2 activity alone or in conjunction with a monomeric form of a peptide having an activity of another B lymphocyte antigen (*e.g.*, B7-1, B7-3) or blocking antibody), prior to transplantation can lead to the binding
20 of the molecule to the natural ligand(s) on the immune cells without transmitting the corresponding costimulatory signal. Blocking B lymphocyte antigen function in this manner prevents cytokine synthesis by immune cells, such as T cells, and thus acts as an immunosuppressant. Moreover, the lack of costimulation may also be sufficient to energize the T cells, thereby inducing tolerance in a subject. Induction of long-term tolerance by B lymphocyte
25 antigen-blocking reagents may avoid the necessity of repeated administration of these blocking reagents. To achieve sufficient immunosuppression or tolerance in a subject, it may also be necessary to block the function of B lymphocyte antigens.

The efficacy of particular blocking reagents in preventing organ transplant rejection or GVHD can be assessed using animal models that are predictive of efficacy in humans. Examples
30 of appropriate systems which can be used include allogeneic cardiac grafts in rats and xenogeneic pancreatic islet cell grafts in mice, both of which have been used to examine the

immunosuppressive effects of CTLA4Ig fusion proteins *in vivo* as described in Lenschow *et al.*, Science 257:789-792 (1992) and Turka *et al.*, Proc Natl Acad Sci USA, 89:11102-11105 (1992). In addition, murine models of GVHD (see Paul ed., FUNDAMENTAL IMMUNOLOGY, Raven Press, New York, 1989, pp. 846-847) can be used to determine the effect of blocking B lymphocyte antigen function *in vivo* on the development of that disease.

Blocking antigen function may also be therapeutically useful for treating autoimmune diseases. Many autoimmune disorders are the result of inappropriate activation of T cells that are reactive against self tissue and which promote the production of cytokines and auto-antibodies involved in the pathology of the diseases. Preventing the activation of autoreactive T cells may reduce or eliminate disease symptoms. Administration of reagents which block costimulation of T cells by disrupting receptor:ligand interactions of B lymphocyte antigens can be used to inhibit T cell activation and prevent production of auto-antibodies or T cell-derived cytokines which may be involved in the disease process. Additionally, blocking reagents may induce antigen-specific tolerance of autoreactive T cells which could lead to long-term relief from the disease. The efficacy of blocking reagents in preventing or alleviating autoimmune disorders can be determined using a number of well-characterized animal models of human autoimmune diseases. Examples include murine experimental autoimmune encephalitis, systemic lupus erythematosus in MRL/lpr/lpr mice or NZB hybrid mice, murine autoimmune collagen arthritis, diabetes mellitus in NOD mice and BB rats, and murine experimental myasthenia gravis (see Paul ed., FUNDAMENTAL IMMUNOLOGY, Raven Press, New York, 1989, pp. 840-856).

Upregulation of an antigen function (preferably a B lymphocyte antigen function), as a means of up regulating immune responses, may also be useful in therapy. Upregulation of immune responses may be in the form of enhancing an existing immune response or eliciting an initial immune response. For example, enhancing an immune response through stimulating B lymphocyte antigen function may be useful in cases of viral infection. In addition, systemic viral diseases such as influenza, the common cold, and encephalitis might be alleviated by the administration of stimulatory forms of B lymphocyte antigens systemically.

Alternatively, anti-viral immune responses may be enhanced in an infected patient by removing T cells from the patient, costimulating the T cells *in vitro* with viral antigen-pulsed APCs either expressing a peptide of the present invention or together with a stimulatory form of a soluble peptide of the present invention and reintroducing the *in vitro* activated T cells into the

patient. Another method of enhancing anti-vital immune responses would be to isolate infected cells from a patient, transfect them with a nucleic acid encoding a protein of the present invention as described herein such that the cells express all or a portion of the protein on their surface, and reintroduce the transfected cells into the patient. The infected cells would now be capable of delivering a costimulatory signal to, and thereby activate, T cells *in vivo*.

In another application, up regulation or enhancement of antigen function (preferably B lymphocyte antigen function) may be useful in the induction of tumor immunity. Tumor cells (e.g., sarcoma, melanoma, lymphoma, leukemia, neuroblastoma, carcinoma) transfected with a nucleic acid encoding at least one peptide of the present invention can be administered to a subject to overcome tumor-specific tolerance in the subject. If desired, the tumor cell can be transfected to express a combination of peptides. For example, tumor cells obtained from a patient can be transfected *ex vivo* with an expression vector directing the expression of a peptide having B7-2-like activity alone, or in conjunction with a peptide having B7-1-like activity and/or B7-3-like activity. The transfected tumor cells are returned to the patient to result in expression of the peptides on the surface of the transfected cell. Alternatively, gene therapy techniques can be used to target a tumor cell for transfection *in vivo*.

The presence of the peptide of the present invention having the activity of a B lymphocyte antigen(s) on the surface of the tumor cell provides the necessary costimulation signal to T cells to induce a T cell mediated immune response against the transfected tumor cells. In addition, tumor cells which lack MHC class I or MHC class II molecules, or which fail to reexpress sufficient amounts of MHC class I or MHC class II molecules, can be transfected with nucleic acid encoding all or a portion of (e.g., a cytoplasmic-domain truncated portion) of an MHC class I α chain protein and β_2 microglobulin protein or an MHC class II α chain protein and an MHC class II β chain protein to thereby express MHC class I or MHC class II proteins on the cell surface. Expression of the appropriate class I or class II MHC in conjunction with a peptide having the activity of a B lymphocyte antigen (e.g., B7-1, B7-2, B7-3) induces a T cell mediated immune response against the transfected tumor cell. Optionally, a gene encoding an antisense construct which blocks expression of an MHC class II associated protein, such as the invariant chain, can also be cotransfected with a DNA encoding a peptide having the activity of a B lymphocyte antigen to promote presentation of tumor associated antigens and induce tumor

specific immunity. Thus, the induction of a T cell mediated immune response in a human subject may be sufficient to overcome tumor-specific tolerance in the subject.

The activity of a protein of the invention may, among other means, be measured by the following methods: Suitable assays for thymocyte or splenocyte cytotoxicity include, without limitation, those described In: CURRENT PROTOCOLS IN IMMUNOLOGY. Coligan *et al.*, eds. Greene Publishing Associates and Wiley-Interscience (Chapter 3, Chapter 7); Herrmann *et al.*, *Proc Natl Acad Sci USA* 78:2488-2492, 1981; Herrmann *et al.*, *J Immunol* 128:1968-1974, 1982; Handa *et al.*, *J Immunol* 135:1564-1572, 1985; Takai *et al.*, *J Immunol* 137:3494-3500, 1986; Takai *et al.*, *J Immunol* 140:508-512, 1988; Herrmann *et al.*, *Proc Natl Acad Sci USA* 78:2488-2492, 1981; Herrmann *et al.*, *J Immunol* 128:1968-1974, 1982; Handa *et al.*, *J Immunol* 135:1564-1572, 1985; Takai *et al.*, *J Immunol* 137:3494-3500, 1986; Bowman *et al.*, *J Virology* 61:1992-1998; Takai *et al.*, *J Immunol* 140:508-512, 1988; Bertagnolli *et al.*, *Cell Immunol* 133:327-341, 1991; Brown *et al.*, *J Immunol* 153:3079-3092, 1994.

Assays for T-cell-dependent immunoglobulin responses and isotype switching (which will identify, among others, proteins that modulate T-cell dependent antibody responses and that affect Th1/Th2 profiles) include, without limitation, those described in: Maliszewski, *J Immunol* 144:3028-3033, 1990; and Mond and Brunswick In: CURRENT PROTOCOLS IN IMMUNOLOGY. Coligan *et al.*, (eds.) Vol 1 pp. 3.8.1-3.8.16, John Wiley and Sons, Toronto 1994.

Mixed lymphocyte reaction (MLR) assays (which will identify, among others, proteins that generate predominantly Th1 and CTL responses) include, without limitation, those described In: CURRENT PROTOCOLS IN IMMUNOLOGY. Coligan *et al.*, eds. Greene Publishing Associates and Wiley-Interscience (Chapter 3, Chapter 7); Takai *et al.*, *J Immunol* 137:3494-3500, 1986; Takai *et al.*, *J Immunol* 140:508-512, 1988; Bertagnolli *et al.*, *J Immunol* 149:3778-3783, 1992.

Dendritic cell-dependent assays (which will identify, among others, proteins expressed by dendritic cells that activate naive T-cells) include, without limitation, those described in: Guery *et al.*, *J Immunol* 134:536-544, 1995; Inaba *et al.*, *J Exp Med* 173:549-559, 1991; Macatonia *et al.*, *J Immunol* 154:5071-5079, 1995; Porgador *et al.*, *J Exp Med* 182:255-260, 1995; Nair *et al.*, *J Virol* 67:4062-4069, 1993; Huang *et al.*, *Science* 264:961-965, 1994; Macatonia *et al.*, *J Exp Med* 169:1255-1264, 1989; Bhardwaj *et al.*, *J Clin Investig* 94:797-807, 1994; and Inaba *et al.*, *J Exp Med* 172:631-640, 1990.

Assays for lymphocyte survival/apoptosis (which will identify, among others, proteins that prevent apoptosis after superantigen induction and proteins that regulate lymphocyte homeostasis) include, without limitation, those described in: Darzynkiewicz *et al.*, *Cytometry* 13:795-808, 1992; Gorczyca *et al.*, *Leukemia* 7:659-670, 1993; Gorczyca *et al.*, *Cancer Res* 53:1945-1951, 1993; Itoh *et al.*, *Cell* 66:233-243, 1991; Zacharchuk, *J Immunol* 145:4037-4045, 1990; Zamai *et al.*, *Cytometry* 14:891-897, 1993; Gorczyca *et al.*, *Internat J Oncol* 1:639-648, 1992.

Assays for proteins that influence early steps of T-cell commitment and development include, without limitation, those described in: Antica *et al.*, *Blood* 84:111-117, 1994; Fine *et al.*, *Cell Immunol* 155: 111-122, 1994; Galy *et al.*, *Blood* 85:2770-2778, 1995; Toki *et al.*, *Proc Nat Acad Sci USA* 88:7548-7551, 1991.

Hematopoiesis Regulating Activity

A GENX protein of the present invention may be useful in regulation of hematopoiesis and, consequently, in the treatment of myeloid or lymphoid cell deficiencies. Even marginal biological activity in support of colony forming cells or of factor-dependent cell lines indicates involvement in regulating hematopoiesis, *e.g.* in supporting the growth and proliferation of erythroid progenitor cells alone or in combination with other cytokines, thereby indicating utility, for example, in treating various anemias or for use in conjunction with irradiation/chemotherapy to stimulate the production of erythroid precursors and/or erythroid cells; in supporting the growth and proliferation of myeloid cells such as granulocytes and monocytes/macrophages (*i.e.*, traditional CSF activity) useful, for example, in conjunction with chemotherapy to prevent or treat consequent myelo-suppression; in supporting the growth and proliferation of megakaryocytes and consequently of platelets thereby allowing prevention or treatment of various platelet disorders such as thrombocytopenia, and generally for use in place of or complimentary to platelet transfusions; and/or in supporting the growth and proliferation of hematopoietic stem cells which are capable of maturing to any and all of the above-mentioned hematopoietic cells and therefore find therapeutic utility in various stem cell disorders (such as those usually treated with transplantation, including, without limitation, aplastic anemia and paroxysmal nocturnal hemoglobinuria), as well as in repopulating the stem cell compartment post irradiation/chemotherapy, either in-vivo or ex-vivo (*i.e.*, in conjunction with bone marrow

transplantation or with peripheral progenitor cell transplantation (homologous or heterologous)) as normal cells or genetically manipulated for gene therapy.

The activity of a protein of the invention may, among other means, be measured by the following methods:

5 Suitable assays for proliferation and differentiation of various hematopoietic lines are cited above.

Assays for embryonic stem cell differentiation (which will identify, among others, proteins that influence embryonic differentiation hematopoiesis) include, without limitation, those described in: Johansson *et al.* *Cellular Biology* 15:141-151, 1995; Keller *et al.*, *Mol. Cell.*
10 *Biol.* 13:473-486, 1993; McClanahan *et al.*, *Blood* 81:2903-2915, 1993.

Assays for stem cell survival and differentiation (which will identify, among others, proteins that regulate lympho-hematopoiesis) include, without limitation, those described in: Methylcellulose colony forming assays, Freshney, In: CULTURE OF HEMATOPOIETIC CELLS. Freshney, *et al.* (eds.) Vol pp. 265-268, Wiley-Liss, Inc., New York, N.Y. 1994; Hirayama *et al.*,
15 *Proc Natl Acad Sci USA* 89:5907-5911, 1992; McNiece and Briddeli, In: CULTURE OF HEMATOPOIETIC CELLS. Freshney, *et al.* (eds.) Vol pp. 23-39, Wiley-Liss, Inc., New York, N.Y. 1994; Neben *et al.*, *Exp Hematol* 22:353-359, 1994; Ploemacher, In: CULTURE OF HEMATOPOIETIC CELLS. Freshney, *et al.* eds. Vol pp. 1-21, Wiley-Liss, Inc., New York, N.Y. 1994; Spoonceret *et al.*, In: CULTURE OF HEMATOPOIETIC CELLS. Freshney, *et al.*, (eds.) Vol pp.
20 163-179, Wiley-Liss, Inc., New York, N.Y. 1994; Sutherland, In: CULTURE OF HEMATOPOIETIC CELLS. Freshney, *et al.*, (eds.) Vol pp. 139-162, Wiley-Liss, Inc., New York, N.Y. 1994.

Tissue Growth Activity

A GENX protein of the present invention also may have utility in compositions used for bone, cartilage, tendon, ligament and/or nerve tissue growth or regeneration, as well as for
25 wound healing and tissue repair and replacement, and in the treatment of burns, incisions and ulcers.

A protein of the present invention, which induces cartilage and/or bone growth in circumstances where bone is not normally formed, has application in the healing of bone fractures and cartilage damage or defects in humans and other animals. Such a preparation
30 employing a protein of the invention may have prophylactic use in closed as well as open fracture reduction and also in the improved fixation of artificial joints. De novo bone formation

induced by an osteogenic agent contributes to the repair of congenital, trauma induced, or oncologic resection induced craniofacial defects, and also is useful in cosmetic plastic surgery.

A protein of this invention may also be used in the treatment of periodontal disease, and in other tooth repair processes. Such agents may provide an environment to attract bone-forming cells, stimulate growth of bone-forming cells or induce differentiation of progenitors of bone-forming cells. A protein of the invention may also be useful in the treatment of osteoporosis or osteoarthritis, such as through stimulation of bone and/or cartilage repair or by blocking inflammation or processes of tissue destruction (collagenase activity, osteoclast activity, etc.) mediated by inflammatory processes.

Another category of tissue regeneration activity that may be attributable to the protein of the present invention is tendon/ligament formation. A protein of the present invention, which induces tendon/ligament-like tissue or other tissue formation in circumstances where such tissue is not normally formed, has application in the healing of tendon or ligament tears, deformities and other tendon or ligament defects in humans and other animals. Such a preparation employing a tendon/ligament-like tissue inducing protein may have prophylactic use in preventing damage to tendon or ligament tissue, as well as use in the improved fixation of tendon or ligament to bone or other tissues, and in repairing defects to tendon or ligament tissue. De novo tendon/ligament-like tissue formation induced by a composition of the present invention contributes to the repair of congenital, trauma induced, or other tendon or ligament defects of other origin, and is also useful in cosmetic plastic surgery for attachment or repair of tendons or ligaments. The compositions of the present invention may provide an environment to attract tendon- or ligament-forming cells, stimulate growth of tendon- or ligament-forming cells, induce differentiation of progenitors of tendon- or ligament-forming cells, or induce growth of tendon/ligament cells or progenitors *ex vivo* for return *in vivo* to effect tissue repair. The compositions of the invention may also be useful in the treatment of tendonitis, carpal tunnel syndrome and other tendon or ligament defects. The compositions may also include an appropriate matrix and/or sequestering agent as a carrier as is well known in the art.

The protein of the present invention may also be useful for proliferation of neural cells and for regeneration of nerve and brain tissue, *i.e.* for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders, which involve degeneration, death or trauma to neural cells or nerve tissue. More specifically, a protein

may be used in the treatment of diseases of the peripheral nervous system, such as peripheral nerve injuries, peripheral neuropathy and localized neuropathies, and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome. Further conditions which may be treated in accordance with the present invention include mechanical and traumatic disorders, such as spinal cord disorders, head trauma and cerebrovascular diseases such as stroke. Peripheral neuropathies resulting from chemotherapy or other medical therapies may also be treatable using a protein of the invention.

Proteins of the invention may also be useful to promote better or faster closure of non-healing wounds, including without limitation pressure ulcers, ulcers associated with vascular insufficiency, surgical and traumatic wounds, and the like.

It is expected that a protein of the present invention may also exhibit activity for generation or regeneration of other tissues, such as organs (including, for example, pancreas, liver, intestine, kidney, skin, endothelium), muscle (smooth, skeletal or cardiac) and vascular (including vascular endothelium) tissue, or for promoting the growth of cells comprising such tissues. Part of the desired effects may be by inhibition or modulation of fibrotic scarring to allow normal tissue to regenerate. A protein of the invention may also exhibit angiogenic activity.

A protein of the present invention may also be useful for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, and conditions resulting from systemic cytokine damage.

A protein of the present invention may also be useful for promoting or inhibiting differentiation of tissues described above from precursor tissues or cells; or for inhibiting the growth of tissues described above.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for tissue generation activity include, without limitation, those described in: International Patent Publication No. WO95/16035 (bone, cartilage, tendon); International Patent Publication No. WO95/05846 (nerve, neuronal); International Patent Publication No. WO91/07491 (skin, endothelium).

Assays for wound healing activity include, without limitation, those described in: Winter, EPIDERMAL WOUND HEALING, pp. 71-112 (Maibach and Rovee, eds.), Year Book Medical

Publishers, Inc., Chicago, as modified by Eaglstein and Menz, *J. Invest. Dermatol* 71:382-84 (1978).

Activin/Inhibin Activity

A GENX protein of the present invention may also exhibit activin- or inhibin-related activities. Inhibins are characterized by their ability to inhibit the release of follicle stimulating hormone (FSH), while activins are characterized by their ability to stimulate the release of follicle stimulating hormone (FSH). Thus, a protein of the present invention, alone or in heterodimers with a member of the inhibin family, may be useful as a contraceptive based on the ability of inhibins to decrease fertility in female mammals and decrease spermatogenesis in male mammals. Administration of sufficient amounts of other inhibins can induce infertility in these mammals. Alternatively, the protein of the invention, as a homodimer or as a heterodimer with other protein subunits of the inhibin-b group, may be useful as a fertility inducing therapeutic, based upon the ability of activin molecules in stimulating FSH release from cells of the anterior pituitary. See, for example, U.S. Pat. No. 4,798,885. A protein of the invention may also be useful for advancement of the onset of fertility in sexually immature mammals, so as to increase the lifetime reproductive performance of domestic animals such as cows, sheep and pigs.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for activin/inhibin activity include, without limitation, those described in: Vale *et al.*, *Endocrinology* 91:562-572, 1972; Ling *et al.*, *Nature* 321:779-782, 1986; Vale *et al.*, *Nature* 321:776-779, 1986; Mason *et al.*, *Nature* 318:659-663, 1985; Forage *et al.*, *Proc Natl Acad Sci USA* 83:3091-3095, 1986.

Chemotactic/Chemokinetic Activity

A protein of the present invention may have chemotactic or chemokinetic activity (*e.g.*, act as a chemokine) for mammalian cells, including, for example, monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells. Chemotactic and chemokinetic proteins can be used to mobilize or attract a desired cell population to a desired site of action. Chemotactic or chemokinetic proteins provide particular advantages in treatment of wounds and other trauma to tissues, as well as in treatment of localized infections. For example,

attraction of lymphocytes, monocytes or neutrophils to tumors or sites of infection may result in improved immune responses against the tumor or infecting agent.

A protein or peptide has chemotactic activity for a particular cell population if it can stimulate, directly or indirectly, the directed orientation or movement of such cell population. Preferably, the protein or peptide has the ability to directly stimulate directed movement of cells. Whether a particular protein has chemotactic activity for a population of cells can be readily determined by employing such protein or peptide in any known assay for cell chemotaxis.

The activity of a protein of the invention may, among other means, be measured by following methods:

Assays for chemotactic activity (which will identify proteins that induce or prevent chemotaxis) consist of assays that measure the ability of a protein to induce the migration of cells across a membrane as well as the ability of a protein to induce the adhesion of one cell population to another cell population. Suitable assays for movement and adhesion include, without limitation, those described in: CURRENT PROTOCOLS IN IMMUNOLOGY, Coligan *et al.*, eds. (Chapter 6.12, MEASUREMENT OF ALPHA AND BETA CHEMOKINES 6.12.1-6.12.28); Taub *et al. J Clin Invest* 95:1370-1376, 1995; Lind *et al. APMIS* 103:140-146, 1995; Muller *et al., Eur J Immunol* 25: 1744-1748; Gruberet *et al. J Immunol* 152:5860-5867, 1994; Johnston *et al., J Immunol* 153: 1762-1768, 1994.

Hemostatic and Thrombolytic Activity

A protein of the invention may also exhibit hemostatic or thrombolytic activity. As a result, such a protein is expected to be useful in treatment of various coagulation disorders (including hereditary disorders, such as hemophilias) or to enhance coagulation and other hemostatic events in treating wounds resulting from trauma, surgery or other causes. A protein of the invention may also be useful for dissolving or inhibiting formation of thromboses and for treatment and prevention of conditions resulting therefrom (such as, for example, infarction of cardiac and central nervous system vessels (e.g., stroke).

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assay for hemostatic and thrombolytic activity include, without limitation, those described in: Linet *et al., J. Clin. Pharmacol.* 26:131-140, 1986; Burdick *et al., Thrombosis Res.*

45:413-419, 1987; Humphrey *et al.*, *Fibrinolysis* 5:71-79 (1991); Schaub, *Prostaglandins* 35:467-474, 1988.

Receptor/Ligand Activity

A protein of the present invention may also demonstrate activity as receptors, receptor
5 ligands or inhibitors or agonists of receptor/ligand interactions. Examples of such receptors and
ligands include, without limitation, cytokine receptors and their ligands, receptor kinases and
their ligands, receptor phosphatases and their ligands, receptors involved in cell—cell
interactions and their ligands (including without limitation, cellular adhesion molecules (such as
10 selectins, integrins and their ligands) and receptor/ligand pairs involved in antigen presentation,
antigen recognition and development of cellular and humoral immune responses). Receptors and
ligands are also useful for screening of potential peptide or small molecule inhibitors of the
relevant receptor/ligand interaction. A protein of the present invention (including, without
limitation, fragments of receptors and ligands) may themselves be useful as inhibitors of
receptor/ligand interactions.

15 The activity of a protein of the invention may, among other means, be measured by the
following methods:

Suitable assays for receptor-ligand activity include without limitation those described in:
CURRENT PROTOCOLS IN IMMUNOLOGY, Ed by Coligan, *et al.*, Greene Publishing Associates and
Wiley-Interscience (Chapter 7.28, Measurement of Cellular Adhesion under static conditions
20 7.28.1-7.28.22), Takai *et al.*, *Proc Natl Acad Sci USA* 84:6864-6868, 1987; Bierer *et al.*, *J. Exp.*
Med. 168:1145-1156, 1988; Rosenstein *et al.*, *J. Exp. Med.* 169:149-160 1989; Stoltenborg *et al.*,
J Immunol Methods 175:59-68, 1994; Stitt *et al.*, *Cell* 80:661-670, 1995.

Anti-Inflammatory Activity

Proteins of the present invention may also exhibit anti-inflammatory activity. The
25 anti-inflammatory activity may be achieved by providing a stimulus to cells involved in the
inflammatory response, by inhibiting or promoting cell—cell interactions (such as, for example,
cell adhesion), by inhibiting or promoting chemotaxis of cells involved in the inflammatory
process, inhibiting or promoting cell extravasation, or by stimulating or suppressing production
of other factors which more directly inhibit or promote an inflammatory response. Proteins
30 exhibiting such activities can be used to treat inflammatory conditions including chronic or acute

conditions), including without limitation inflammation associated with infection (such as septic shock, sepsis or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine-induced lung injury, inflammatory bowel disease, Crohn's disease or resulting from over production of cytokines such as TNF or IL-1. Proteins of the invention may also be useful to treat anaphylaxis and hypersensitivity to an antigenic substance or material.

Tumor Inhibition Activity

In addition to the activities described above for immunological treatment or prevention of tumors, a protein of the invention may exhibit other anti-tumor activities. A protein may inhibit tumor growth directly or indirectly (such as, for example, via ADCC). A protein may exhibit its tumor inhibitory activity by acting on tumor tissue or tumor precursor tissue, by inhibiting formation of tissues necessary to support tumor growth (such as, for example, by inhibiting angiogenesis), by causing production of other factors, agents or cell types which inhibit tumor growth, or by suppressing, eliminating or inhibiting factors, agents or cell types which promote tumor growth.

Other Activities

A protein of the invention may also exhibit one or more of the following additional activities or effects: inhibiting the growth, infection or function of, or killing, infectious agents, including, without limitation, bacteria, viruses, fungi and other parasites; effecting (suppressing or enhancing) bodily characteristics, including, without limitation, height, weight, hair color, eye color, skin, fat to lean ratio or other tissue pigmentation, or organ or body part size or shape (such as, for example, breast augmentation or diminution, change in bone form or shape); effecting biorhythms or circadian cycles or rhythms; effecting the fertility of male or female subjects; effecting the metabolism, catabolism, anabolism, processing, utilization, storage or elimination of dietary fat, lipid, protein, carbohydrate, vitamins, minerals, cofactors or other nutritional factors or component(s); effecting behavioral characteristics, including, without limitation, appetite, libido, stress, cognition (including cognitive disorders), depression (including depressive disorders) and violent behaviors; providing analgesic effects or other pain reducing effects; promoting differentiation and growth of embryonic stem cells in lineages other than hematopoietic lineages; hormonal or endocrine activity; in the case of enzymes, correcting

deficiencies of the enzyme and treating deficiency-related diseases; treatment of hyperproliferative disorders (such as, for example, psoriasis); immunoglobulin-like activity (such as, for example, the ability to bind antigens or complement); and the ability to act as an antigen in a vaccine composition to raise an immune response against such protein or another material or entity which is cross-reactive with such protein.

Neural disorders in general include Parkinson's disease, Alzheimer's disease, Huntington's disease, multiple sclerosis, amyotrophic lateral sclerosis (ALS), peripheral neuropathy, tumors of the nervous system, exposure to neurotoxins, acute brain injury, peripheral nerve trauma or injury, and other neuropathies, epilepsy, and/or tremors.

EQUIVALENTS

From the foregoing detailed description of the specific embodiments of the invention, it should be apparent that particular novel compositions and methods involving nucleic acids, polypeptides, antibodies, detection and treatment have been described. Although these particular embodiments have been disclosed herein in detail, this has been done by way of example for purposes of illustration only, and is not intended to be limiting with respect to the scope of the appended claims that follow. In particular, it is contemplated by the inventors that various substitutions, alterations, and modifications may be made as a matter of routine for a person of ordinary skill in the art to the invention without departing from the spirit and scope of the invention as defined by the claims. Indeed, various modifications of the invention in addition to those described herein will become apparent to those skilled in the art from the foregoing description and accompanying figures. Such modifications are intended to fall within the scope of the appended claims.

THIS PAGE INTENTIONALLY LEFT BLANK

THIS PAGE INTENTIONALLY LEFT BLANK

THIS PAGE INTENTIONALLY LEFT BLANK

THIS PAGE INTENTIONALLY LEFT BLANK

THIS PAGE INTENTIONALLY LEFT BLANK

THIS PAGE INTENTIONALLY LEFT BLANK

THIS PAGE INTENTIONALLY LEFT BLANK

THIS PAGE INTENTIONALLY LEFT BLANK

17	26232744 (33, 34)	Novel Protein sim. GBank glt174849p g44391 URE1-JAEIN - UREASE ALPHA SUBUNIT (UREA AMIDOHYDROLASE) Isotop [Arabidopsis thaliana]	Contains protein domain (PF00449) - Urease		264600
18	60246804 (35, 36)	Novel Protein sim. GBank g12281102 (A002333) - SF16			29331827, 264355, 264357, 264638, 264558
19	60276524 (37, 38)	Novel Protein sim. GBank p15506112 p143872 LUP - ECOLI - ABC TRANSPORTER ATP-BINDING PROTEIN LUP	UNCLASSIFIED transport		22278996, 264807, 264910, 264620, 264693 264602
20	20724558 (39, 40)	Novel Protein sim. GBank p15506112 p143872 LUP - ECOLI - ABC TRANSPORTER ATP-BINDING PROTEIN LUP	UNCLASSIFIED		22278996, 264806, 265008, 265010, 265011, 264602, 264605, 264766, 264688, 21906764, 264691, 18108376, 264636, 18108387, 264485
21	80417554 (41, 42)	Novel Protein sim. GBank glt1730203 p15042 GATM - RAT - GLYCINE AMIDINOTRANSFERASE PRECURSOR (L- ARGININE GLYCINE AMIDINOTRANSFERASE) (TRANSAMIDINASE) (AT)			264685
22	11705658 (43, 44)	Novel Protein sim. GBank glt187732 p15 CAB07071 - (Z82771) lade25 [Mycobacterium tuberculosis]	Contains protein domain (PF00441) - Acyl-CoA dehydrogenase		264886, 264907, 294009, 264600, 264602, 264603, 264605, 264682, 264766, 32833986, 264636, 264496
23	80418176 (45, 46)	Novel Protein sim. GBank glt187732 p15 CAB07071 - (Z82771) lade25 [Mycobacterium tuberculosis]	dehydrogenase		264600
24	26291697 (47, 48)	Novel Protein sim. GBank glt187732 p15 CAB07071 - (Z82771) lade25 [Mycobacterium tuberculosis]			264593
25	80253774 (49, 50)	Novel Protein sim. GBank glt187732 p15 CAB07071 - (Z82771) lade25 [Mycobacterium tuberculosis]	UNCLASSIFIED ribosomal		22278996, 58182435, 265010, 264566
26	80253994 (51, 52)	Novel Protein sim. GBank glt187732 p15 CAB07071 - (Z82771) lade25 [Mycobacterium tuberculosis]	UNCLASSIFIED ribosomal		18108370, 39689423, 264635, 264555
27	80253995 (53, 54)	Novel Protein sim. GBank glt187732 p15 CAB07071 - (Z82771) lade25 [Mycobacterium tuberculosis]	UNCLASSIFIED ribosomal		264593
28	79485561 (55, 56)	Novel Protein sim. GBank glt187732 p15 CAB07071 - (Z82771) lade25 [Mycobacterium tuberculosis]	UNCLASSIFIED kinase		264638
29	82446755 (57, 58)	Novel Protein sim. GBank glt187732 p15 CAB07071 - (Z82771) lade25 [Mycobacterium tuberculosis]	UNCLASSIFIED kinase		264601, 264762, 264766, 264769, 264636
30	79198333 (59, 60)	Novel Protein sim. GBank glt187732 p15 CAB07071 - (Z82771) lade25 [Mycobacterium tuberculosis]	UNCLASSIFIED		264908, 265019, 264887, 21906764, 21906766
31	19846158 (61, 62)	Novel Protein sim. GBank g13560504 (AF027770) - unknown [Mycobacterium smegmatis]	UNCLASSIFIED		264534
32	82449495 (63, 64)	Novel Protein sim. GBank g13560504 (AF027770) - unknown [Mycobacterium smegmatis]	UNCLASSIFIED		264905, 264605, 264762, 264766, 264687, 264689
33	79592628 (65, 66)	Novel Protein sim. GBank g12159030 p1 G54507 - hypothetical protein M1865 - Methanococcus jannaschii	UNCLASSIFIED		264687
34	81467657 (67, 68)	Novel Protein sim. GBank g12159030 p1 G54507 - hypothetical protein M1865 - Methanococcus jannaschii	UNCLASSIFIED		60432288, 264600, 264602, 264760, 18108357, 264769, 265020, 264891
35	92005170 (69, 70)	Novel Protein sim. GBank g12159030 p1 G54507 - hypothetical protein M1865 - Methanococcus jannaschii	UNCLASSIFIED		264600, 264887, 264636, 264639
36	19842042 (71, 72)	Novel Protein sim. GBank g12159030 p1 G54507 - hypothetical protein M1865 - Methanococcus jannaschii	synthase		264566
37	20369215 (73, 74)	Novel Protein sim. GBank g12159030 p1 G54507 - hypothetical protein M1865 - Methanococcus jannaschii	dehydrogenase		264603

56	g1227508 (115, 116)	Novel Protein sim. GBank g10516074gipAAC45918-1pAF06194 - (AF061943) protease- derived STE20 like kinase Psk [Homo sapiens]	Contains protein domain (PF00069) - Eukaryotic protein kinase domain	Kinase	5612375, 264259, 60432046, 35696052, 66712502, 264959, 265000, 265010, 265011, 264681, 29148784, 35695917, 60170615, 264691, 264692, 264693, 16103374, 258502, 26182323, 60432113 264600, 264689, 264636
59	80077371 (117, 118)	Novel Protein sim. GBank g11729208pP46930RPE_MYCLE - PUTATIVE UNDECAPRENYL-PHOSPHATE ALPHAN- ACETYL-GLUCOSAMINYL-TRANSFERASE	Contains protein domain (PF00953) - Glycosyl transferase	transferase	
60	12969341 (119, 120)	Novel Protein sim. GBank g11729208pP46930RPE_MYCLE - PUTATIVE UNDECAPRENYL-PHOSPHATE ALPHAN- ACETYL-GLUCOSAMINYL-TRANSFERASE			
61	80426606 (121, 122)	Novel Protein sim. GBank g11710216 (J78260) - unknown [Homo sapiens]		glycoprotein	264689 264766
62	13504666 (123, 124)				
63	16474553 (125, 126)	Novel Protein sim. GBank g11729208pP46930RPE_MYCLE - PUTATIVE UNDECAPRENYL-PHOSPHATE ALPHAN- ACETYL-GLUCOSAMINYL-TRANSFERASE			
64	20724578 (127, 128)	Novel Protein sim. GBank g1420545gipJ47041 - transportase homolog (insertion element SAEL) - Alcaligenes eutrophus		UNCLASSIFIED	264630
65	79326308 (129, 130)	Novel Protein sim. GBank g15123128pJ06134KPYV_MYCTU - PYRUVATE KINASE (PK)	Contains protein domain (PF00224) - Pyruvate kinase	Kinase	264563
66	45854384 (131, 132)	Novel Protein sim. GBank g13920723hmsJCA42219 - ALDOLASE (ALD) putative ABC transporter [Stigmatomycetes codicicola]		transport	22278995, 264558
67	76952543 (133, 134)	Novel Protein sim. GBank g15123128pJ06134KPYV_MYCTU - ALANINE DEHYDROGENASE (40 KD ANITGEN)		dehydrogenase	265021
68	76817382 (135, 136)				
69	75841764 (137, 138)				
70	76871328 (139, 140)				
71	65897456 (141, 142)				
72	87734877 (143, 144)	Novel Protein sim. GBank g1441582gipJAD20157 - (A0006282) unknown protein [Arabidopsis thaliana]			
73	60025241 (145, 146)				
74	20372410 (147, 148)	Novel Protein sim. GBank g1285108gipJCA160141 - (AL021787) vesicular protein sorting [Schistosoma mansoni]			
75	11816032 (149, 150)				
76	95105303 (151, 152)	Novel Protein sim. GBank g14466811hmsJCA382121 - (AL035601) putative protein [Arabidopsis thaliana]			
77	10144718 (153, 154)	Novel Protein sim. GBank g1854065gipJCA583371 - (X63413) UBB [Human herpesvirus 6]			
78	8758258 (155, 156)				

79	[941401157, 158] Novel Protein sim. GBank g 5689453 b b BA330.0.1 - (A02029381) KIAA1059 protein (Homo sapiens)	Contains protein domain PH domain	35696386, 22278989, 23531822, 29331824, 23331824, 26712502, 2646008, 2646008, 2646008, 68712502, 2646008, 2646008, 265009, 2646008, 60170831, 5512038, 33109964, 265017, 265018, 264288, 264769, 56181902, 21004755, 21004769, 29148794, 265020, 264630, 264631, 264632, 264633, 60431528, 35598423, 264631, 264632, 294634, 264636, 264639, 83373044, 264564, 294566, 264457
80	[92314540 (155, 160)]	UNCLASSIFIED	294769, 264601, 255006, 264910, 264604, 264605, 264634, 264635, 264605, 264762, 264637, 264592, 264628, 264907, 264891, 264508, 264567, 264909, 264766, 264605
81	[20467247 (161, 162)]	UNCLASSIFIED	294769, 264601, 255006, 264910, 264604, 264605, 264634, 264635, 264605, 264762, 264637, 264592, 264628, 264907, 264891, 264508, 264567, 264909, 264766, 264605
82	[16331386 (163, 164)] Novel Protein sim. GBank g 17234426 b Q10258 YD2A, SCHPO, HYPOTHETICAL 59.0 KD PROTEIN C56F3.10 (INCHROMOSOME1) methylmalonate semi-aldehyde dehydrogenase (Oryza saliva)	reductase	264567
83	[94741180 (165, 165)] Novel Protein sim. GBank g 3420273 (AC004697) - unknown protein (Arabidopsis thaliana)	dehydrogenase	264567
84	[80355372 (167, 168)] Novel Protein sim. GBank g 17336456 b P45306 SAT1, RAT - SULFATE ANION TRANSPORTER 1 (CANALICULAR SULFATE TRANSPORTER) [SULFATE/CARBONATE ANTI-PORTER] Novel Protein sim. GBank g 276996 b P1570082 - phosphatidylcholine transferase (Homo sapiens) Novel Protein sim. GBank g 326026 b P1570082 - (A021897) hypothetical protein Rv1112 (Mycobacterium tuberculosis)	transport	264488, 264508, 264509, 264905, 264908, 264908, 264511, 264591, 264593, 264594, 264595, 264596, 264756, 264603, 264780, 264631, 18108351, 264762, 264602, 264764, 264684, 264766, 264686, 264632, 264637, 264557, 264638, 264639, 18108355, 264566, 264508, 264606, 264907, 264908, 264909, 264910, 264760, 264763, 264764, 264766, 264769, 264769, 35695955, 264636, 264637, 264605, 264762, 264687, 264769, 18108374, 264635, 264688, 264910
85	[80499600 (168, 170)] Novel Protein sim. GBank g 276996 b P1570082 - phosphatidylcholine transferase (Homo sapiens) Novel Protein sim. GBank g 326026 b P1570082 - (A021897) hypothetical protein Rv1112 (Mycobacterium tuberculosis)	transferase	264488, 264508, 264509, 264905, 264908, 264908, 264511, 264591, 264593, 264594, 264595, 264596, 264756, 264603, 264780, 264631, 18108351, 264762, 264602, 264764, 264684, 264766, 264686, 264632, 264637, 264557, 264638, 264639, 18108355, 264566, 264508, 264606, 264907, 264908, 264909, 264910, 264760, 264763, 264764, 264766, 264769, 264769, 35695955, 264636, 264637, 264605, 264762, 264687, 264769, 18108374, 264635, 264688, 264910
86	[39558043 (171, 172)] Novel Protein sim. GBank g 326026 b P1570082 - (A021897) hypothetical protein Rv1112 (Mycobacterium tuberculosis)	UNCLASSIFIED	264488, 264508, 264509, 264905, 264908, 264908, 264511, 264591, 264593, 264594, 264595, 264596, 264756, 264603, 264780, 264631, 18108351, 264762, 264602, 264764, 264684, 264766, 264686, 264632, 264637, 264557, 264638, 264639, 18108355, 264566, 264508, 264606, 264907, 264908, 264909, 264910, 264760, 264763, 264764, 264766, 264769, 264769, 35695955, 264636, 264637, 264605, 264762, 264687, 264769, 18108374, 264635, 264688, 264910
87	[13558005 (173, 174)]	UNCLASSIFIED	264488, 264508, 264509, 264905, 264908, 264908, 264511, 264591, 264593, 264594, 264595, 264596, 264756, 264603, 264780, 264631, 18108351, 264762, 264602, 264764, 264684, 264766, 264686, 264632, 264637, 264557, 264638, 264639, 18108355, 264566, 264508, 264606, 264907, 264908, 264909, 264910, 264760, 264763, 264764, 264766, 264769, 264769, 35695955, 264636, 264637, 264605, 264762, 264687, 264769, 18108374, 264635, 264688, 264910

98	94322125 (191, 192)	Novel Protein sim. GBank g1459509[gb]BAA15602.1] - (A9023175) KIA00558 protein [Homo sapiens]	UNCLASSIFIED	2277895, 2277899, 264590, 2633192, 2331823, 2559092, 2514493, 264590, 264906, 264907, 264908, 264909, 265007, 265008, 264910, 265009, 264930, 265010, 265017, 264604, 265019, 18103351, 264288, 264766, 264768, 264769, 21900765, 21900767, 21905769, 265020, 264632, 33657182, 35655763, 264628, 264629, 18108378, 264631, 264636, 18108361, 264559, 18108382, 83373044, 22775002, 264508
99	70605200 (193, 194)	Novel Protein sim. GBank g14583359[emb]CAB40388.1] - (A005285) OxyR [Escherichia coli]	UNCLASSIFIED	264909
100	78427000 (195, 196)	Novel Protein sim. GBank g1101693[gb]BAA10430] - (D5402) hypothetical protein [Synchocystis sp.]	UNCLASSIFIED	264605
101	20465324 (197, 198)	Novel Protein sim. GBank g1101693[gb]BAA10430] - (D5402) hypothetical protein [Synchocystis sp.]	UNCLASSIFIED	264605
102	78404113 (199, 200)	Novel Protein sim. GBank g1480387[em]IS37445 - gene	UNCLASSIFIED	264633
103	80202268 (201, 202)	Novel Protein sim. GBank g1480387[em]IS37445 - gene	UNCLASSIFIED	265020, 264102, 263972
104	20467259 (203, 204)	Novel Protein sim. GBank g12894166[emb]CAA11773.1] - (A223985) PCZA361.18 [Amycolatopsis orientalis]	UNCLASSIFIED	264605
105	20466358 (205, 206)	Novel Protein sim. GBank g1731040[gb]P54509[OH]_BACSU - HYPOTHETICAL HELICASE IN SINI-GOVT INTERGENIC REGION	UNCLASSIFIED	264605
106	80247572 (207, 208)	Novel Protein sim. GBank g1854065[emb]CAA58337] - (X83413) U88 [Human herpesvirus 6]	UNCLASSIFIED	264591, 264592, 264602
107	79605206 (209, 210)	Novel Protein sim. GBank g11685117 [U0770] - [unreviewed]	UNCLASSIFIED	264508
108	26382058 (211, 212)	Crossophila melanogaster	complement	264511, 265009
109	80057911 (213, 214)	Novel Protein sim. GBank g1705959[gb]P54729[B54]_MCUSE - B54 PROTEIN	UNCLASSIFIED	29331824, 264591, 21900754, 265019
110	80237656 (215, 216)	Novel Protein sim. GBank g148722[gb]PAC032244.1JAF-15075 - (AF150755) [Bacillus subtilis]	ATPase, associated	18108374, 26501567, 22776896, 264113, 264906, 264907, 264908, 265017, 234810, 264909, 264930, 264766, 264637
111	95194148 (217, 218)	Novel Protein sim. GBank g12330791[emb]CAB11255] - (Z98601) carboxypeptidase precursor	UNCLASSIFIED	264766, 264636, 264637
112	75655823 (219, 220)	Schistosoma mansoni	transport	264637
113	35965458 (221, 222)			264554
114	79656038 (223, 224)		UNCLASSIFIED	264908
115	17959438 (225, 226)		UNCLASSIFIED	265007
116	80502101 (227, 228)		UNCLASSIFIED	264769

115	80251003 (226, 230)	Novel Protein sim. GBank g12244532 (U93872) - ORF 73, contains large complex repeat CR 73 [Kaposi's sarcoma-associated herpesvirus]	UNCLASSIFIED	15845156, 12545900, 33556970, 264592, 21904754, 27460264, 18108379, 36595423, 264535, 52544332, 18108382, 264905, 264906, 264907, 264908, 264909, 264910, 264758, 265510, 264763, 264682, 264764, 264756, 264685, 264686, 264768, 264769, 33557023, 264693, 33557109, 264828, 18108374, 264631, 264632, 264634, 264635, 264637, 264638, 264639, 56526488, 264555, 264586
116	81266869 (231, 232)		UNCLASSIFIED	263974
117	79635695 (231, 234)		UNCLASSIFIED	263974
118	80222170 (233, 236)		UNCLASSIFIED	263974
119	91013071 (237, 238)	Novel Protein sim. GBank g1732526 (U22307) - alpha2(I) collagen [Caenorhabditis elegans]	UNCLASSIFIED	22278596, 29331824, 5043328, 265007, 66433438, 264603, 264605, 18108351, 264769, 264689, 265020, 264534, 27489281, 264558, 83373044, 18108385, 264564
120	87564611 (239, 240)	Novel Protein sim. GBank g1213191[S9150157] - cyclin-dependent kinase chain SRB10 - yeast [Saccharomyces cerevisiae]	kinase	264603
121	80026153 (241, 242)			264605
122	20457620 (243, 244)	Novel Protein sim. GBank g12052147[embCAB08137] - (ZNF452) ksgA [Mycobacterium tuberculosis]	Contains protein domain (PF00308) - Ribosomal RNA, adenosine dimethylases	264605
123	8758278 (245, 246)	Novel Protein sim. GBank g12833385[g14314][UGST_SORBI - GRANULE-BOUND GLYCOSYL (STARCH) SYNTHASE PRECURSOR	synthase	264604
124	79104071 (247, 248)	Novel Protein sim. GBank g1475542 (U82955) - glutamate receptor delta-4 subunit [Rattus norvegicus]	-misc_channel	18108394, 18108397, 2655038, 2656007, 2656010, 2656011, 18108355, 18108379, 18108380, 18108384, 264936, 264936, 265009, 264956, 22279002
125	87797996 (249, 250)	Novel Protein sim. GBank g15102785[embCAB43200.1] - (AL079308) putative transcriptional regulator [Streptomyces coelicolor]	Ligand-gated ion channel	264511
126	56701243 (251, 252)		UNCLASSIFIED	264605
127	20467267 (253, 254)	Novel Protein sim. GBank g11301210[gp173320PHOB_PSEAE_PHCSPHATE	Contains protein domain (PF00072) - Response regulator receiver domain	264605
128	80244473 (255, 256)	Novel Protein sim. GBank g12606493[pp129300]GCB_ECOLI - HYPOTHEICAL 100.5 KD PROTEIN IN IMP-CY5 INTERGENIC REGION	UNCLASSIFIED	264605
129	93290543 (257, 258)	Novel Protein sim. GBank g12606493[pp129300]GCB_ECOLI - HYPOTHEICAL 100.5 KD PROTEIN IN IMP-CY5 INTERGENIC REGION	UNCLASSIFIED	264605
130	90051583 (259, 260)	Novel Protein sim. GBank g1476538[gp1554860] - ABC transporter P4C4 chain - Mycobacterium tuberculosis	UNCLASSIFIED	264605
131	94990022 (261, 263)	Novel Protein sim. GBank g1476538[gp1554860] - ABC transporter P4C4 chain - Mycobacterium tuberculosis	transport	264605

132	10887692 (263, 264)	Novel Protein sim. GBank_gli1677340(emb)[CA807068] - (Z32771) acsA3 [Mycobacterium tuberculosis]		carboxylase	264638 264603, 264669
133	9433063 (263, 266)	Novel Protein sim. GBank_gli495835(emb)[CA14932.1] - (AL049630) putative NADH dehydrogenase [Streptomyces coelicolor]		dehydrogenase	264605, 264605, 265021
134	78834860 (267, 268)	Novel Protein sim. GBank_gli1460074(emb)[CA807049] - (Z77250) hypothetical protein Rv2566 [Mycobacterium tuberculosis]			264634
135	19885057 (269, 270)	Novel Protein sim. GBank_gli125896(emb)[CA72511] - (Y13070) [oligopolyglutamate synthase [Streptomyces coelicolor]		synthase	264609
136	78846003 (271, 272)	Novel Protein sim. GBank_gli520387(emb)[CA846781] - (A4243439) porphobiliprotein [Lactinaria major]		UNCLASSIFIED	264683, 264685, 264686, 264691, 264692, 264693
137	78619770 (273, 274)	Novel Protein sim. GBank_gli568937(emb)[CA152073.1] - (AL089732) putative nuclease [Streptomyces coelicolor CA32]		dehydrogenase	2227898, 265007, 264910, 03433356, 245010, 264602, 264605, 267168, 264688, 264764, 264693, 32333986, 16108314, 264698
140	78625759 (273, 280)	Novel Protein sim. GBank_gli375756(emb)[CA21315] - (AL031545) nuS family DNA mismatch repair protein [Schizosaccharomyces pombe]		UNCLASSIFIED	264609
141	207700394 (281, 282)	Novel Protein sim. GBank_gli3334172(esp)[P9303BAC2_MOUSE - TRANSCRIPTION REGULATOR PROTEIN BACH2 (B1B AND CNC HOMOLOG 2)		nuclease	264602, 265017
142	80028104 (283, 284)	Novel Protein sim. GBank_gli406074(emb)[Y436] - BURGE - INSERTION ELEMENT S402 HYPOTHETICAL 24 KD PROTEIN		UNCLASSIFIED	264600
143	11072274 (285, 286)	Novel Protein sim. GBank_gli2622760(emb)[CA016683] - (AL021646) uncD2 [Mycobacterium tuberculosis]		UNCLASSIFIED	263078, 264620, 264910, 264632, 264408, 264563, 264584, 264591, 264556, 264408, 264623, 264639
144	95009102 (287, 288)	Novel Protein sim. GBank_gli2016347(esp)[CA31585] - (AL031863) 1-evidence-predicted by content, 1-method-genefinder:084; 1-method_score=46.31; 1-evidence and [Drosophila melanogaster]		UNCLASSIFIED	2227898, 264602
145	80027058 (289, 290)	Novel Protein sim. GBank_gli2016347(esp)[CA31585] - (AL021989) hypothetical protein Rv0366 [Mycobacterium tuberculosis]		UNCLASSIFIED	264687
146	13085662 (291, 292)	Novel Protein sim. GBank_gli2016347(esp)[CA31585] - (AL021989) hypothetical protein Rv0366 [Mycobacterium tuberculosis]		UNCLASSIFIED	264687
147	94320368 (293, 294)	Novel Protein sim. GBank_gli2016347(esp)[CA31585] - (AL021989) hypothetical protein Rv0366 [Mycobacterium tuberculosis]		UNCLASSIFIED	264687
148	80248604 (295, 296)	Novel Protein sim. GBank_gli2016347(esp)[CA31585] - (AL021989) hypothetical protein Rv0366 [Mycobacterium tuberculosis]		UNCLASSIFIED	264687

149	60249373 (297, 298)	Novel Protein sim. GBank glt17230733hp011040Y081 MYCTU - HYPOTHETICAL ABC TRANSPORTER ATP-BINDING PROTEIN CY56.01	Contains protein domain (PF00009) - ABC transporter	transport	285010, 284600, 284601, 284603, 284604, 27483855, 284636
150	20254748 (280, 300)	Novel Protein sim. GBank glt3724125smj[CA11005] - (AJ224340) maltosephosphorylase [Lactobacillus sarlifancoc]			284600
151	20726398 (301, 302)	Novel Protein sim. GBank glt7293125hp0765 [PEOB_ECO1] - PHOSPHOENOLPYRUVATE KINASE (PF00009)	Contains protein domain (PF01678) - Metalloenzyme superfamily	UNCLASSIFIED	284602
152	95002877 (303, 304)	Novel Protein sim. GBank glt497325hp05867 [14TM_RHISN - HYPOTHETICAL HYDROLASE/PEPTIDASE Y41M]		peptidase	284602
153	80256655 (305, 306)	Novel Protein sim. GBank glt1249306hp0701 [EUKIE - VITELLOGENIN I PRECURSOR (VTG) CONTAINS A PROTEINELLIN 1 (LV1), PHOSVITIN (PV), LIPOTYLLIN 2 (LV2)]			284953
154	82305968 (307, 308)	Novel Protein sim. GBank glt411669 [pJN0443 - transcription initiation factor sigma homolog htrB - Streptomyces aureofaciens]	Contains protein domain (PF00140) - Sigma-70 factor	UNCLASSIFIED	284602, 284652, 284651, 284654 284605
155	20429859 (309, 310)	Novel Protein sim. GBank glt628710hp [j541739 - hypothetical protein - Escherichia coli]		UNCLASSIFIED	284595
156	35564742 (311, 312)	Novel Protein sim. GBank glt3695013 [AF052586] - CtrA [Pseudomonas aeruginosa]	Contains protein domain (PF00742) - 4Fe-4S iron sulfur cluster binding proteins, NifH-HcX family	hydrolyase	284091
157	10354887 (313, 314)	Novel Protein sim. GBank glt1073072hp [C55543 - cnaU protein - Pseudomonas syringae pv. syringae]		UNCLASSIFIED	284905
158	79761036 (315, 316)	Novel Protein sim. GBank glt171024hp049789 [RL30_3 - STRCO - 50S RIBOSOMAL PROTEIN L30]	Contains protein domain (PF00400) - WD domain G-beta repeat	UNCLASSIFIED	285008 284605
159	76860378 (317, 318)	Novel Protein sim. GBank glt171024hp049789 [RL30_3 - STRCO - 50S RIBOSOMAL PROTEIN L30]	Contains protein domain (PF00400) - WD domain G-beta repeat	UNCLASSIFIED	285008 284605
160	11073119 (319, 320)	Novel Protein sim. GBank glt171024hp049789 [RL30_3 - STRCO - 50S RIBOSOMAL PROTEIN L30]	Contains protein domain (PF00400) - WD domain G-beta repeat	UNCLASSIFIED	285008 284605
161	80055007 (321, 322)	Novel Protein sim. GBank glt171024hp049789 [RL30_3 - STRCO - 50S RIBOSOMAL PROTEIN L30]	Contains protein domain (PF00400) - WD domain G-beta repeat	UNCLASSIFIED	285008 284605
162	80016371 (323, 324)	Novel Protein sim. GBank glt530486 [pmsC] A946028.31 - (AL031685) dJ863C23.2 (novel protein [Pomo sapensis])	Contains protein domain (PF00027) - Ribosomal protein L30p17e	ribosomal prot	22278096, 284600, 284603, 356595917, 32833986, 35695423, 284636
163	11692306 (325, 326)	Novel Protein sim. GBank glt265168 [terp] [CA115705] - (AL009204) putative protease [Streptomyces coelicolor]	Contains protein domain (PF00097) - Zinc finger, C2H2 type (RING finger)	antileukin	284112, 284532, 22279002
164	80077602 (327, 328)	Novel Protein sim. GBank glt4116478hp [AA020378] - (AF125959) transposase [Mycobacterium avium]		UNCLASSIFIED	284639 284905, 284607, 284600
165	10850507 (329, 330)	Novel Protein sim. GBank glt4116478hp [AA020378] - (AF125959) transposase [Mycobacterium avium]		UNCLASSIFIED	284691
166	88955003 (331, 332)	Novel Protein sim. GBank glt265168 [terp] [CA115705] - (AL009204) putative protease [Streptomyces coelicolor]		UNCLASSIFIED	284600, 284486
167	16395460 (333, 334)	Novel Protein sim. GBank glt4116478hp [AA020378] - (AF125959) transposase [Mycobacterium avium]		UNCLASSIFIED	285010
168	80079362 (335, 336)	Novel Protein sim. GBank glt76177 [p] [QCECF - hypothetical 38 kD protein (fla 2' region) - Escherichia coli]		UNCLASSIFIED	284600
169	80239581 (337, 338)	Novel Protein sim. GBank glt76177 [p] [QCECF - hypothetical 38 kD protein (fla 2' region) - Escherichia coli]			284556, 284457, 284558, 284559

170	7561264 (335, 340)	Novel Protein sim. GBank glt40886ipf27847ygc_ECOLI - HYPOTHETICAL 15.4 KD PROTEIN IN REQ-PLD INTERGENIC REGION (P138)	Contains protein domain (PF01810) - LysE type translocator	264906 264956, 264604
171	9523073 (341, 342)	Novel Protein sim. GBank glt4210959gltAA02048.11 - AGG (Stromatolium melioid)	transport	264769
172	37379707 (343, 344)	Novel Protein sim. GBank glt32854epf02387RL2_ECOLI - 50S RIBOSOMAL PROTEIN L2	Contains protein domain (PF00528) - binding-protein-dependent transport systems inner membrane component	264769
173	57529660 (345, 346)	Novel Protein sim. GBank glt2072722gmhCA08326 - (Z95121) manA (Mycobacterium tuberculosis)	Contains protein domain (PF00181) - Ribosomal Proteins L2	264769
174	9523078 (347, 348)	Novel Protein sim. GBank glt005178 (U00187) - similar to PRLC41724 chicken LD (limb deformity) gene product and IC (containing also P-rich region similar to collagen subtilisin)	transport	264510, 264593, 264602, 264603, 264605, 264762, 264693
175	79756270 (349, 350)	Novel Protein sim. GBank glt2072722gmhCA08326 - (Z95121) manA (Mycobacterium tuberculosis)	isomerase	264565
176	80066896 (351, 352)	Novel Protein sim. GBank glt005178 (U00187) - similar to PRLC41724 chicken LD (limb deformity) gene product and IC (containing also P-rich region similar to collagen subtilisin)	UNCLASSIFIED	264907, 264910, 264981, 264558
177	86844852 (353, 354)	Novel Protein sim. GBank glt232673ajmncAB109521 - (Z98268) hypothetical protein Rv1865 (Mycobacterium tuberculosis)	Contains protein domain (PF01513) - Domain of unknown function	264768, 8043419, 264637, 264638, 264769, 20331825, 60432249, 18108370, 264689, 18108387, 32833986, 22778993, 265020, 264600, 264601, 264602, 264603, 264604, 264605, 264635, 264762, 264636, 264606, 264564, 264637, 264638, 264486, 60433358, 264765
178	75559526 (355, 356)	Novel Protein sim. GBank glt19055096 (U18187) - kinesin-73 (Drosophila melanogaster)	struct	264693, 33657109, 264635
179	20263112 (357, 358)	Novel Protein sim. GBank glt169367epfP45256IDNAE - REP. LATIVE DNA HELICASE	UNCLASSIFIED	264593
180	80489538 (359, 360)	Novel Protein sim. GBank glt3170615 (A7059485) - DOC4 (Mus musculus)	helicase	264769
181	75553569 (361, 362)	Novel Protein sim. GBank glt1076627epfIS54172 - inorganic pyrophosphatase (EC 3.6.1.1) - common tobacco	UNCLASSIFIED	21906767, 264635, 264639, 18108394
182	80577899 (363, 364)	Novel Protein sim. GBank glt4371280jpbAA0181381 - (AC008260) hypothetical protein (Arabidopsis thaliana)	UNCLASSIFIED	264259, 35696052, 66184345, 264571, 260318, 33657109, 264525, 264556, 264690
183	11614017 (365, 366)	Novel Protein sim. GBank glt4371280jpbAA0181381 - (AC008260) hypothetical protein (Arabidopsis thaliana)	UNCLASSIFIED	264610
184	10174167 (367, 368)	Novel Protein sim. GBank glt4371280jpbAA0181381 - (AC008260) hypothetical protein (Arabidopsis thaliana)	UNCLASSIFIED	

185	21603022 (369, 370)	Novel Protein sim. GBank g 306617 emb CAA1838.1 - (AL023304) putative trnaa transport regulator [Schistosomacromyces pombe]		UNCLASSIFIED	264604
186	80070329 (371, 372)	Novel Protein sim. GBank g 2828802 sp P94409 CYCLE_BACSU - HYPOTHETICAL 53.3 KD PROTEIN IN SFP-GERKA INTERGENIC REGION		transport	264555
187	80186511 (373, 374)	Novel Protein sim. GBank g 3150260 emb CAA1879.1 - (AL023604) cyclin [Schistosomacromyces pombe]		UNCLASSIFIED	264359
188	20464942 (375, 376)	Novel Protein sim. GBank g 245853 pf S172938 - ntix protein - Mycobacterium leprae		kinase	264605
189	8233215 (377, 378)			UNCLASSIFIED	35698052, 264602, 264605, 264762, 264689, 35690917, 18108370, 18108372, 264638, 264555
190	80068821 (379, 380)	Novel Protein sim. GBank g 1181244 db BAA1927.1 - (AF001480) SIMILAR TO PYRUVATE OXIDASE AND PYRUVATE KINASE [Bacillus subtilis]		synthase	264563
191	86095012 (381, 382)	Novel Protein sim. GBank g 120286 sp P27248 PKRP_STROCH - FK506-BINDING PROTEIN (PEPTIDYL-PROLYL CIS-TRANS ISOMERASE) (PPIASE) (ROTAMASE)		isomerase	264508, 264604, 264605, 264789, 264555
192	16333379 (383, 384)				264567
193	79910127 (385, 386)				264608, 264593
194	20464949 (387, 388)				264605
195	13518389 (389, 390)			transport	264636
196	90055569 (391, 392)	Novel Protein sim. GBank g 4980682 gb AAQ03547.1 AE00171 - (AE001718) ABC transporter, ATP-binding protein [Thermotoga maritima]			264600, 264689, 264638
197	80248685 (393, 394)	Novel Protein sim. GBank g 1705461 pf P36969 BIOXA_ERWHE - ADENYCYL METHIONINE-L-AMINO-7- OXONONANOATE AMINOTRANSFERASE (7,8-DIAMINO PELARGONIC ACID AMINOTRANSFERASE) (DAPA AMINOTRANSFERASE)		gaba	
198	79163535 (395, 396)	Novel Protein sim. GBank g 1123209 sp D2777 HIERP_SCHYMA_6 - PHOSPHOFRUCTOKINASE (PHOSPHOFRUCTOKINASE) (PHOSPHO-HEXOKINASE)		kinase	264602, 264682, 264682, 18108374
199	78880715 (397, 398)	Novel Protein sim. GBank g 1161203 emb CA006110 - (Z03359) gnd [Mycobacterium tuberculosis]		Phosphofructokinase	
200	78413849 (399, 400)	Novel Protein sim. GBank g 2642222 AF020085 - telomere-associated recQ-like helicase [Lathyrus maydis]		Continuing protein domain (PF00392) - β -phosphoketone dehydrogenases	264636, 265008
201	86949324 (401, 402)	Novel Protein sim. GBank g 2864379 emb CAA17481.1 - (Y14573) ring finger protein [Hordeum vulgare]		UNCLASSIFIED	264595, 264596
				UNCLASSIFIED	79331826, 255207, 264512, 33657402, 264596, 265017, 18108351, 264682, 264683, 264767, 264628, 55810764, 264634, 264635, 59182323, 69432113, 22275000

202	7958046 (403, 404)	Novel Protein sim. GBank g1231772ip10089jCHS1_USTMA - CHITIN SYNTHASE 1,3-CHITIN UDP ACETYL-GLUCOSAMINYL TRANSFERASE 1	Contains protein domain (PF01844) - synthase Chitin synthase	264600
203	79843927 (405, 406)	Novel Protein sim. GBank g1150424ip10089jBAA13220 - (G68964) similar to yeast adenylate cyclase (S56716) (Homo sapiens)		22278995, 29331822, 29331825, 29331827, 264606, 21906754, 264683, 21906766, 21906769, 25694423, 264556
204	79855186 (407, 408)	Novel Protein sim. GBank g126336ip10089jCAB13310 - (Z3911) similar to hypothetical proteins [Bacillus subtilis]	UNCLASSIFIED	264609
205	10090543 (409, 410)	Novel Protein sim. GBank g126336ip10089jCAB13310 - (Z3911) similar to hypothetical proteins [Bacillus subtilis]	transport	264609
206	8758472 (411, 412)	Novel Protein sim. GBank g1234381ip10089j - polybromo 1 protein - chicken	UNCLASSIFIED	264604
207	20754522 (413, 414)	Novel Protein sim. GBank g1234381ip10089j - polybromo 1 protein - chicken	UNCLASSIFIED	264556
208	20289281 (415, 416)	Novel Protein sim. GBank g1250104ip10089jHISYP - MYCTU - PROLYL-TRNA SYNTHETASE (PROLINE--TRNA LIGASE) (PROKS)		264605
209	80071069 (417, 418)	Novel Protein sim. GBank g1250104ip10089jHISYP - MYCTU - PROLYL-TRNA SYNTHETASE (PROLINE--TRNA LIGASE) (PROKS)		264605, 264689
210	80168800 (419, 420)	Novel Protein sim. GBank g1250104ip10089jHISYP - MYCTU - PROLYL-TRNA SYNTHETASE (PROLINE--TRNA LIGASE) (PROKS)		264605, 264907, 264939, 264786, 264687, 264691, 264625, 18108374, 264638
211	80024539 (421, 422)	Novel Protein sim. GBank g103160ip10089jIP - 2065536 - IP(SLR - Immunoglobulin heavy chain constant region repeat		263978
212	83442474 (423, 424)	Novel Protein sim. GBank g103160ip10089jIP - 2065536 - IP(SLR - Immunoglobulin heavy chain constant region repeat	UNCLASSIFIED	264508, 264905, 264906, 264907, 264908, 264909, 264782, 264534, 264632, 264634, 264535, 264536, 264488
213	80249522 (425, 426)	Novel Protein sim. GBank g1122359ip10089j3123LEU1 - MYCLE - 3 - (ISOPROPYL MALATE DEHYDRATASE LARGE SUBUNIT (ISOMERASE) (IPMI)	Contains protein domain (PF00330) - isomerase family (conazole hydrolase)	22278995, 264308, 264800, 264602, 264603, 264605, 33637023, 264585, 264486
214	80079381 (427, 428)	Novel Protein sim. GBank g1116238ip10089j19421CH90 - COXRU - 60 KD CHAPERONIN (PROTEIN CPN60) (GROEL PROTEIN) (HEAT SHOCK PROTEIN 8)	Contains protein domain (PF00119) - TCP-1cpn60 chaperonin family	264600, 264693
215	14373283 (429, 430)	Novel Protein sim. GBank g1116238ip10089j19421CH90 - COXRU - 60 KD CHAPERONIN (PROTEIN CPN60) (GROEL PROTEIN) (HEAT SHOCK PROTEIN 8)	eph	
216	80177718 (431, 432)	Unknown gene product [Homo sapiens]	UNCLASSIFIED	264629
217	79803634 (433, 434)	Novel Protein sim. GBank g112348ip10089j49754VP41 - HUMAN - VACUOLAR PROTEIN VPS41 HOMOLOG (S53)	Zinc finger C2H2 type	264448
218	80258475 (435, 436)	Novel Protein sim. GBank g1172348ip10089j49754VP41 - HUMAN - VACUOLAR PROTEIN VPS41 HOMOLOG (S53)	264508	
219	204336797 (437, 438)	Novel Protein sim. GBank g1172348ip10089j49754VP41 - HUMAN - VACUOLAR PROTEIN VPS41 HOMOLOG (S53)	rimpolymerase	264594
220	13495972 (439, 440)	Novel Protein sim. GBank g1172348ip10089j49754VP41 - HUMAN - VACUOLAR PROTEIN VPS41 HOMOLOG (S53)	synthase	264604
221	11287498 (441, 442)	Novel Protein sim. GBank g1172348ip10089j49754VP41 - HUMAN - VACUOLAR PROTEIN VPS41 HOMOLOG (S53)	nucleosidase	264689
		Novel Protein sim. GBank g1172348ip10089j49754VP41 - HUMAN - VACUOLAR PROTEIN VPS41 HOMOLOG (S53)	UNCLASSIFIED	264555

222	7865202 (443, 444)	Novel Protein sim. GBank gl187758[emjC4307049] - (282770) hypothetical protein Rv01432 [Mycobacterium tuberculosis]	UNCLASSIFIED	264605, 264768, 265084-23
223	8305369 (445, 446)		UNCLASSIFIED	264906, 264907, 264903
224	78557820 (447, 448)		UNCLASSIFIED	264904, 264993
225	78559541 (449, 450)	Novel Protein sim. GBank gl227485 [hbnBAA42151s] - (D41159) 3-2 gene product [Homo sapiens]	UNCLASSIFIED	264982
226	79172397 (451, 452)	Novel Protein sim. GBank gl888245 (U29468) - C95C10.7 gene product [Caenorhabditis elegans]	UNCLASSIFIED	22278968, 264112, 33657023, 263981
227	81777196 (453, 454)		UNCLASSIFIED	35635917, 264636, 264907
228	79872285 (455, 456)			264768, 264907, 264908, 264632, 264593, 264639
229	79832686 (457, 458)			264906, 264910
230	11013209 (459, 460)		UNCLASSIFIED	264631
231	26522207 (461, 462)	Novel Protein sim. GBank gl183514[emjC4471733] - (Y10744) homoserine O-acetyltransferase [Leptospira meyeri]	UNCLASSIFIED	264906, 264603, 264603, 264632
232	80055035 (463, 464)	Novel Protein sim. GBank gl2642340 (AF032970) - (Z63398) Putative proteinase [Pseudomonas putida]	UNCLASSIFIED	264604
233	80053054 (465, 466)	Novel Protein sim. GBank gl3510205 (AF030883) - polypeptidase [Fugu subsp.]	UNCLASSIFIED	264389
234	7523988 (467, 468)	Novel Protein sim. GBank gl2104609[emjC4308805] - (Z63398) PdcA [Mycobacterium leprae]	UNCLASSIFIED	264106, 264905
235	80203671 (469, 470)			264681
236	76940001 (471, 472)	Novel Protein sim. GBank gl3875920[emjC4304111] - (Z61843) predicted using GeneHancer; similar to collagen; (Z61843) EMBL060-400 comes from the gene cDNA EST E811; D60880 comes from the gene [Caenorhabditis elegans]	UNCLASSIFIED	264905, 265010, 264603, 264762, 264682, 264636, 264638, 264486
237	11755273 (473, 474)	Novel Protein sim. GBank	transport	264259, 264769
238	79461401 (475, 476)	Novel Protein sim. GBank		
239	82435190 (477, 478)	Novel Protein sim. GBank	UNCLASSIFIED	264681
240	21635575 (479, 480)	Novel Protein sim. GBank	UNCLASSIFIED	264681
241	60377307 (481, 482)	Novel Protein sim. GBank gl3875920[emjC4304111] - (Z61843) predicted using GeneHancer; similar to collagen; (Z61843) EMBL060-400 comes from the gene cDNA EST E811; D60880 comes from the gene [Caenorhabditis elegans]	UNCLASSIFIED	264905, 264909, 264769, 264639
242	82148454 (483, 484)		UNCLASSIFIED	264486, 264907, 264908, 264511, 264760, 264764, 264892, 264635, 264637
243	79633207 (485, 486)	Novel Protein sim. GBank gl2624302[emjC4415075] - (AL008967) ad [Mycobacterium tuberculosis]	UNCLASSIFIED	264905
244	80248562 (487, 488)	Novel Protein sim. GBank gl2920625 (AF044489) - vprE protein [Escherichia coli]	UNCLASSIFIED	264602, 264602, 264605, 264769, 264689
245	79663543 (489, 490)	Novel Protein sim. GBank gl5420387[emjC466719.1] - (AJ243459) proteolipophagyron [Leishmania major]	UNCLASSIFIED	264607, 264758
246	79165629 (491, 492)			264637, 18108381, 18100387, 264655

247	79873185 (494, 494)	Novel Protein sim. GBank g[11859056]emj[C400648] - (Z5592) argp [Mycobacterium tuberculosis]	kinase	264909, 264991, 35696423, 18106387
248	80488983 (495, 496)	Novel Protein sim. GBank g[11865749]p4240q[1P8_CORGL - ATP SYNTHASE BETA CHAIN	synthase	35959286, 254307, 264511, 264602, 264766, 264688, 355021, 35959385, 18100385
249	76764645 (497, 498)	Novel Protein sim. GBank g[11865749]p4240q[1P8_CORGL - ATP SYNTHASE BETA CHAIN	UNCLASSIFIED	264007, 264910, 265011, 264792, 264538
250	79515960 (499, 500)	Novel Protein sim. GBank g[11865749]p4240q[1P8_CORGL - ATP SYNTHASE BETA CHAIN	UNCLASSIFIED	264007, 264910, 265011, 264792, 264538
251	84359489 (501, 502)	Novel Protein sim. GBank g[11865749]p4240q[1P8_CORGL - ATP SYNTHASE BETA CHAIN	UNCLASSIFIED	264007, 264910, 265011, 264792, 264538
252	79737756 (503, 504)	Novel Protein sim. GBank g[3327165]p[BA31651] - (A8014576) KIAA0675 protein (Homo sapiens)	UNCLASSIFIED	265018, 21995765, 21995768, 265020, 27486261, 27486265, 35959763, 18108376, 264558, 264559, 264565
253	20443124 (505, 506)	Novel Protein sim. GBank g[3303680]emj[C4A18513] - (AL022374) putative ATP-dependent DNA helicase	helicase	264685, 264687, 264632
254	8027421 (507, 508)	Novel Protein sim. GBank g[3303680]emj[C4A18513] - (AL022374) putative ATP-dependent DNA helicase	helicase	264685, 264687, 264632
255	11398315 (509, 510)	Novel Protein sim. GBank g[3303680]emj[C4A18513] - (AL022374) putative ATP-dependent DNA helicase	UNCLASSIFIED	264508, 264509, 294602, 264687, 265021, 264486
256	80028138 (511, 512)	Novel Protein sim. GBank g[3303680]emj[C4A18513] - (AL022374) putative ATP-dependent DNA helicase	UNCLASSIFIED	264508, 264509, 294602, 264687, 265021, 264486
257	20289282 (513, 514)	Novel Protein sim. GBank g[3303680]emj[C4A18513] - (AL022374) putative ATP-dependent DNA helicase	UNCLASSIFIED	264508, 264509, 294602, 264687, 265021, 264486
258	20459464 (515, 516)	Novel Protein sim. GBank g[3303680]emj[C4A18513] - (AL022374) putative ATP-dependent DNA helicase	UNCLASSIFIED	264508, 264509, 294602, 264687, 265021, 264486
259	79910152 (517, 518)	Novel Protein sim. GBank g[3303680]emj[C4A18513] - (AL022374) putative ATP-dependent DNA helicase	UNCLASSIFIED	264508, 264509, 294602, 264687, 265021, 264486
260	20375437 (519, 520)	Novel Protein sim. GBank g[3303680]emj[C4A18513] - (AL022374) putative ATP-dependent DNA helicase	UNCLASSIFIED	264508, 264509, 294602, 264687, 265021, 264486
261	20285983 (521, 522)	Novel Protein sim. GBank g[3303680]emj[C4A18513] - (AL022374) putative ATP-dependent DNA helicase	UNCLASSIFIED	264508, 264509, 294602, 264687, 265021, 264486
262	80180317 (523, 524)	Novel Protein sim. GBank g[3303680]emj[C4A18513] - (AL022374) putative ATP-dependent DNA helicase	UNCLASSIFIED	264508, 264509, 294602, 264687, 265021, 264486
263	80059045 (525, 526)	Novel Protein sim. GBank g[3303680]emj[C4A18513] - (AL022374) putative ATP-dependent DNA helicase	UNCLASSIFIED	264508, 264509, 294602, 264687, 265021, 264486
264	87370826 (527, 528)	Novel Protein sim. GBank g[3303734]p[BA23531] - (A8011177) KIAA0605 protein (Homo sapiens)	protease	264259, 264908, 21906754, 265018, 265019, 265020

2665	65355646 (529, 530)	Novel Protein sm. GBank	glt4556624b[JB316834.1] - (A03023207) KIA0099a protein [Homo sapiens]	kinase	254483, 35696286, 70331924, 56182181, 36969032, 284508, 764903, 264908, 264907, 66712502, 264592, 264908, 264511, 264512, 254910, 264592, 264595, 264758, 264596, 55811386, 264600, 265017, 264603, 264604, 254605, 264760, 18108351, 264762, 264681, 254764, 264288, 264766, 264768, 264769, 21906765, 21906767, 21906769, 265020, 264691, 33657023, 33657106, 33657182, 264628, 35696423, 35656855, 264630, 264631, 264632, 264634, 264635, 264636, 264555, 264638, 83373044, 56526486, 71683518, 264564, 264566, 264486
2666	79589075 (531, 532)	Novel Protein sm. GBank	glt507627[pefNP_003182.1] pARS - ribosomal RNA synthetase	UNCLASSIFIED	254600
2667	11362222 (533, 534)	Novel Protein sm. GBank	glt507636[pJA005333.1] pAE01070 - (AE01070) glucose-phosphate isomerase [Thermotoga maritima]	UNCLASSIFIED	254628
2668	76700566 (535, 536)	Novel Protein sm. GBank	glt168274[ppP44698NKD.0] pHEIN - PROBABLE P-NUCLEOTIDASE PRECURSOR	UNCLASSIFIED	254602
2669	80023810 (537, 538)	Novel Protein sm. GBank	glt168274[ppP44698NKD.0] pHEIN - PROBABLE P-NUCLEOTIDASE PRECURSOR	UNCLASSIFIED	254693
2670	84381144 (539, 540)	Novel Protein sm. GBank	glt168274[ppP44698NKD.0] pHEIN - PROBABLE P-NUCLEOTIDASE PRECURSOR	UNCLASSIFIED	254698
2671	79552931 (541, 542)	Novel Protein sm. GBank	glt507636[pJA005333.1] pAE01070 - (AE01070) glucose-phosphate isomerase [Thermotoga maritima]	UNCLASSIFIED	254608, 264653
2672	9514718 (543, 544)	Novel Protein sm. GBank	glt507636[pJA005333.1] pAE01070 - (AE01070) glucose-phosphate isomerase [Thermotoga maritima]	synthase	264605
2673	12840694 (545, 546)	Novel Protein sm. GBank	glt168274[ppP44698NKD.0] pHEIN - PROBABLE P-NUCLEOTIDASE PRECURSOR	UNCLASSIFIED	264698
2674	39524245 (547, 548)	Novel Protein sm. GBank	glt32323159 [AF005335] - translation initiation factor eIF2C [Oryctolagus cuniculus]	UNCLASSIFIED	264654
2675	82787041 (548, 550)	Novel Protein sm. GBank	glt34920[ppP21997TSSGP.1] VOLCA - SULFATED SURFACE GLYCOPROTEIN 185 (SSG 185)	UNCLASSIFIED	264697, 264908, 264909, 264766, 264768, 264691, 264632, 264638
2676	86671073 (551, 552)	Novel Protein sm. GBank	glt34920[ppP21997TSSGP.1] VOLCA - SULFATED SURFACE GLYCOPROTEIN 185 (SSG 185)	UNCLASSIFIED	265008, 60432229
2677	80078735 (553, 554)	Novel Protein sm. GBank	glt123021[ppP20984DBG.0] BACSU - SPOOB-ASSOCIATED GTP-BINDING PROTEIN	ribosomal protein	264600, 18108397
2678	12566947 (555, 556)	Novel Protein sm. GBank	glt78833[ppP303812.1] bvbB protein - Micrococcus luteus	UNCLASSIFIED	264689
2679	95292719 (557, 558)	Novel Protein sm. GBank	glt78833[ppP303812.1] bvbB protein - Micrococcus luteus	nuclease	264608, 264604, 21900764, 264638, 264557, 264404
2680	5603617 (559, 560)	Novel Protein sm. GBank	glt31231[ppPQ1986ATLW2.0] CAEL - HYPOTHETICAL 46.2 KD TRP-ASP REPEATS CONTAINING PROTEIN D2013.2 IN CHROMOSOME II	UNCLASSIFIED	264759
2681	80249599 (561, 562)	Novel Protein sm. GBank	glt31231[ppPQ1986ATLW2.0] CAEL - HYPOTHETICAL 46.2 KD TRP-ASP REPEATS CONTAINING PROTEIN D2013.2 IN CHROMOSOME II	UNCLASSIFIED	18108392, 264634, 264555, 264556, 264557, 264558
2682	15398632 (563, 564)	Novel Protein sm. GBank	glt31231[ppPQ1986ATLW2.0] CAEL - HYPOTHETICAL 46.2 KD TRP-ASP REPEATS CONTAINING PROTEIN D2013.2 IN CHROMOSOME II	UNCLASSIFIED	265019
2683	70514231 (565, 566)	Novel Protein sm. GBank	glt31231[ppPQ1986ATLW2.0] CAEL - HYPOTHETICAL 46.2 KD TRP-ASP REPEATS CONTAINING PROTEIN D2013.2 IN CHROMOSOME II	UNCLASSIFIED	264535

284	91212160 (561, 566)	Novel Protein sim. GBank g12429094 (U58632) - acetyl kylan esterase; AxaA [Thermotoga neopolitana]	Contains protein domain (PF00300) - Phosphoglycerate mutase family	UNCLASSIFIED	35060052, 23331828, 264508, 264905, 264600, 264602, 264605, 264682, 264764, 264765, 21905794, 18108376, 264636, 264559, 18105387
285	8157940 (569, 570)	Novel Protein sim. GBank g1207254 (emb1C4300305) - (Z59120) rHE [Mycobacterium tuberculosis]	Contains protein domain (PF00270) - DEAD/DEAF box helicase	UNCLASSIFIED	264603
287	12745521 (573, 574)	Novel Protein sim. GBank g1763323 (pbst157676 - (S74439) silk fibroin heavy chain (C-terminal) [Bombyx mori] mor1-silkworms, Peptide Partial, 633 aa) [Bombyx mori]		UNCLASSIFIED	35060052, 264768, 264638
288	20756502 (575, 576)	Novel Protein sim. GBank g11870060 (emb1C4300600) - (Z62338) hypothetical protein rv1019 [Mycobacterium tuberculosis]		UNCLASSIFIED	264689 264557
289	60043604 (577, 578)	Novel Protein sim. GBank g11870060 (emb1C4300600) - (Z62338) hypothetical protein rv1019 [Mycobacterium tuberculosis]	Contains protein domain (PF00440) - Bacterial regulatory proteins, tetR family	UNCLASSIFIED	264593, 264600
290	80430175 (579, 580)	Novel Protein sim. GBank g1207254 (emb1C4300305) - TRG-RIML INTERGENIC REGION PRECURSOR		UNCLASSIFIED	264768
291	20747431 (581, 582)	Novel Protein sim. GBank g1625182 (L38019) - mitochondrial glutaryl-tRNA synthetase [Saccharomyces cerevisiae]		UNCLASSIFIED	264601
292	80052555 (583, 584)	Novel Protein sim. GBank g11870060 (emb1C4300600) - (Z62338) hypothetical protein rv1019 [Mycobacterium tuberculosis]		UNCLASSIFIED	264605
293	80052519 (585, 586)	Novel Protein sim. GBank g11718065 (pPSU82) (UVRD_MYCLE - PUTATIVE DNA HELICASE II HOMOLOG		UNCLASSIFIED	264909, 264605, 264687, 264689, 264692
294	78630303 (587, 588)	g117422 (sp10040) (CRB_DROME - CRUIGS PROTEIN PRECURSOR (95F)	Contains protein domain (PF00009) - EGF-like domain	UNCLASSIFIED	35060052, 264606, 265011, 264628, 55811576
295	79444180 (589, 590)	Novel Protein sim. GBank g1181819 (dbj BAAT1595 - (D82384) a variant of TSC-22 [Gallus gallus]		UNCLASSIFIED	52644507, 23331822, 264652, 265020, 264639
296	79607076 (591, 592)	Novel Protein sim. GBank g13549789 (dbj BAJ3403 - (AB012226) SecA [Vibrio alginolyticus]		UNCLASSIFIED	264508
297	79631257 (593, 594)	Novel Protein sim. GBank g15689387 (emb1C4302004.1) - (AL109653) putative membrane protein [Streptomyces collabor A3(2)]		UNCLASSIFIED	264905, 264687, 264638
298	80418898 (595, 596)	Novel Protein sim. GBank g12429094 (U58632) - acetyl kylan esterase; AxaA [Thermotoga neopolitana]		UNCLASSIFIED	264905, 264691, 264639, 264766

314	07645112 (627, 628)	Novel Protein sm. GBank gji3681583 (AF092175) - Iaros (Dano ren)	Contains protein domain (PF00302) - GATA zinc finger	264259, 264432289, 26331028, 264805, 264406, 264908, 264909, 265008, 264910, 6043229, 33857402, 60433436, 33109594, 265011, 265017, 264603, 265016, 264286, 264610, 264615, 264616, 264617, 264618, 601705, 264619, 264620, 264621, 264622, 264623, 264624, 264625, 264626, 264627, 264628, 264629, 264630, 264631, 264632, 264633, 264634, 264635, 264636, 264637, 264638, 264639, 264640, 264641, 264642, 264643, 264644, 264645, 264646, 264647, 264648, 264649, 264650, 264651, 264652, 264653, 264654, 264655, 264656, 264657, 264658, 264659, 264660, 264661, 264662, 264663, 264664, 264665, 264666, 264667, 264668, 264669, 264670, 264671, 264672, 264673, 264674, 264675, 264676, 264677, 264678, 264679, 264680, 264681, 264682, 264683, 264684, 264685, 264686, 264687, 264688, 264689, 264690, 264691, 264692, 264693, 264694, 264695, 264696, 264697, 264698, 264699, 264700, 264701, 264702, 264703, 264704, 264705, 264706, 264707, 264708, 264709, 264710, 264711, 264712, 264713, 264714, 264715, 264716, 264717, 264718, 264719, 264720, 264721, 264722, 264723, 264724, 264725, 264726, 264727, 264728, 264729, 264730, 264731, 264732, 264733, 264734, 264735, 264736, 264737, 264738, 264739, 264740, 264741, 264742, 264743, 264744, 264745, 264746, 264747, 264748, 264749, 264750, 264751, 264752, 264753, 264754, 264755, 264756, 264757, 264758, 264759, 264760, 264761, 264762, 264763, 264764, 264765, 264766, 264767, 264768, 264769, 264770, 264771, 264772, 264773, 264774, 264775, 264776, 264777, 264778, 264779, 264780, 264781, 264782, 264783, 264784, 264785, 264786, 264787, 264788, 264789, 264790, 264791, 264792, 264793, 264794, 264795, 264796, 264797, 264798, 264799, 264800, 264801, 264802, 264803, 264804, 264805, 264806, 264807, 264808, 264809, 264810, 264811, 264812, 264813, 264814, 264815, 264816, 264817, 264818, 264819, 264820, 264821, 264822, 264823, 264824, 264825, 264826, 264827, 264828, 264829, 264830, 264831, 264832, 264833, 264834, 264835, 264836, 264837, 264838, 264839, 264840, 264841, 264842, 264843, 264844, 264845, 264846, 264847, 264848, 264849, 264850, 264851, 264852, 264853, 264854, 264855, 264856, 264857, 264858, 264859, 264860, 264861, 264862, 264863, 264864, 264865, 264866, 264867, 264868, 264869, 264870, 264871, 264872, 264873, 264874, 264875, 264876, 264877, 264878, 264879, 264880, 264881, 264882, 264883, 264884, 264885, 264886, 264887, 264888, 264889, 264890, 264891, 264892, 264893, 264894, 264895, 264896, 264897, 264898, 264899, 264900, 264901, 264902, 264903, 264904, 264905, 264906, 264907, 264908, 264909, 264910, 264911, 264912, 264913, 264914, 264915, 264916, 264917, 264918, 264919, 264920, 264921, 264922, 264923, 264924, 264925, 264926, 264927, 264928, 264929, 264930, 264931, 264932, 264933, 264934, 264935, 264936, 264937, 264938, 264939, 264940, 264941, 264942, 264943, 264944, 264945, 264946, 264947, 264948, 264949, 264950, 264951, 264952, 264953, 264954, 264955, 264956, 264957, 264958, 264959, 264960, 264961, 264962, 264963, 264964, 264965, 264966, 264967, 264968, 264969, 264970, 264971, 264972, 264973, 264974, 264975, 264976, 264977, 264978, 264979, 264980, 264981, 264982, 264983, 264984, 264985, 264986, 264987, 264988, 264989, 264990, 264991, 264992, 264993, 264994, 264995, 264996, 264997, 264998, 264999, 265000, 265001, 265002, 265003, 265004, 265005, 265006, 265007, 265008, 265009, 265010, 265011, 265012, 265013, 265014, 265015, 265016, 265017, 265018, 265019, 265020, 265021, 265022, 265023, 265024, 265025, 265026, 265027, 265028, 265029, 265030, 265031, 265032, 265033, 265034, 265035, 265036, 265037, 265038, 265039, 265040, 265041, 265042, 265043, 265044, 265045, 265046, 265047, 265048, 265049, 265050, 265051, 265052, 265053, 265054, 265055, 265056, 265057, 265058, 265059, 265060, 265061, 265062, 265063, 265064, 265065, 265066, 265067, 265068, 265069, 265070, 265071, 265072, 265073, 265074, 265075, 265076, 265077, 265078, 265079, 26508
-----	---------------------	---	--	--

347	75158135 (883, 894)	Novel Protein sim. GBank g1350p38759THAL YEAST - HYPOTHETICAL PRTS 100% ID. NOVEL ERG1-NMDC INTERGENIC REGION Novel Protein sim. GBank g1073010g101507612 - upb protein - Escherichia coli	UNCLASSIFIED	255006, 255008, 265010, 265018, 265367, 263981
348	80020208 (685, 696)	Novel Protein sim. GBank g1073010g101507612 - upb protein - Escherichia coli	transport	264602, 18106351, 18106387
349	17282112 (697, 698)	Novel Protein sim. GBank g132616591emb1CAB009171 - (Z77137) hypothetical protein Pw1277 (Mycobacterium tuberculosis)	nuclease	265007
350	80502370 (699, 700)	Novel Protein sim. GBank g12659357jmb1CXA17921 - (AL022117) hypothetical protein [Schistosoma mansoni pombe]	glycoprotein	264769, 264905, 264908
351	80501805 (701, 702)	Novel Protein sim. GBank g14418302g1AA02030071 - (AF108716) copia type pol polyprotein [Zea mays]	protease	264595
352	11611585 (703, 704)	Novel Protein sim. GBank SUBUNIT [UREA AMIDOHYDROLASE]	UNCLASSIFIED	264604
353	80061653 (705, 706)	Novel Protein sim. GBank SUBUNIT [UREA AMIDOHYDROLASE]	UNCLASSIFIED	264628
354	56826130 (707, 708)	Novel Protein sim. GBank FACTORS PUTATIVE POSITIVE TRANSCRIPTION REGULATOR BVCA	UNCLASSIFIED	264609, 264595, 264683, 22279002
355	80046344 (709, 710)	Novel Protein sim. GBank g1497637 (J03833) - cytochrome oxidase d subunit I [Escherichia coli]	transcription factor	264909, 264591, 264582
356	80043835 (711, 712)	Novel Protein sim. GBank g12290930 (AF006000) - Big1 [Bordetella pertussis]	oxidase	264605
357	86070566 (713, 714)	Novel Protein sim. GBank g13510639 (AF049344) - UOP- GALNAc:polypeptide N-acetylglucosaminyltransferase T5 [Rattus norvegicus]	UNCLASSIFIED	264768
358	37032756 (715, 716)	Novel Protein sim. GBank GLUCANOTRANSFERASE	UNCLASSIFIED	264604, 264769
359	90501488 (717, 718)	Novel Protein sim. GBank GLUCANOTRANSFERASE	UNCLASSIFIED	264594
360	80025748 (719, 720)	Novel Protein sim. GBank GLUCANOTRANSFERASE	transferase	22278966, 264250, 20331822, 20331824, 264605, 58511987, 265022
361	80584075 (721, 722)	Novel Protein sim. GBank GLUCANOTRANSFERASE	amylase	264688
362	13083465 (723, 724)	Novel Protein sim. GBank GLUCANOTRANSFERASE	UNCLASSIFIED	264604, 264769
363	78750145 (725, 726)	Novel Protein sim. GBank GLUCANOTRANSFERASE	UNCLASSIFIED	264594
364	82443593 (727, 728)	Novel Protein sim. GBank DEHYDROGENASE [CHAM K (NADH)UBIQUINONE OXIDOREDUCTASE CHAIN 11] (NUO11)	dehydrogenase	264769, 264602, 264604, 264509, 264762, 264638, 264486

[illegible]

376	8605153 (755, 789)	Novel Protein sim. GBank gl1076013(pj14A9930 - carB	Contains protein domain (PF00689) -	UNCLASSIFIED	265503, 264555
379	8050437 (757, 785)	protein homolog - Mycobacterium bovis (strain H37Rv)	Carbamoyl-phosphate synthase (CPSase)	synthase	264769
380	8006093 (759, 790)	Novel Protein sim. GBank gl12558(pj14A02174) -	Contains protein domain (PF01011) -	dehydrogenase	264604
381	11759027 (761, 762)	(C1267) glucose dehydrogenase (Escherichia coli)	POD enzyme repeat	UNCLASSIFIED	264584
382	8005437 (763, 764)	Novel Protein sim. GBank gl3327156(pj14A031836) -			264592
383	83256025 (765, 766)	(AB014561) KIAA0661 [protein (Homo sapiens)]			264595, 265017, 265021, 264638, 87768518,
384	90314255 (767, 768)	Novel Protein sim. GBank gl1073456(pj14S47810 - probable	Contains protein domain (PF00465) -	dehydrogenase	22279002
385	10237679 (769, 770)	alcohol dehydrogenase (EC 1.1.1.1) - Escherichia coli	iron-containing alcohol dehydrogenases	UNCLASSIFIED	264259, 26331822, 60432289, 26331827,
386	79839434 (771, 772)	Novel Protein sim. GBank gl1460074(pj14A01049) -	Contains protein domain (PF01841) -	UNCLASSIFIED	264288, 264768, 263967, 65274791,
387	11796037 (773, 774)	(Z77260) hypothetical protein Rv2586 [Mycobacterium tuberculosis]	Transglutaminase-like superfamily		35695655, 263981, 8337044, 264567
388	87741378 (775, 776)	Novel Protein sim. GBank gl4240169(pj14A074853.1) -	Contains protein domain (PF00646) -	UNCLASSIFIED	264602
389	76316971 (777, 778)	(A0630647) KIAA0840 protein (Homo sapiens)	F-box domain	UNCLASSIFIED	264600
390	80079949 (779, 780)	Novel Protein sim. GBank gl854065(pj14A058337) -		UNCLASSIFIED	264482
391	7657302 (781, 782)	(X33413) U85 [Human herpesvirus 6]		UNCLASSIFIED	264908
392	79796056 (783, 784)	Novel Protein sim. GBank gl3378523(pj14A08887) -		synthase	264602, 21906764
393	33206031 (785, 786)	(A009932) cyclomaltohexanase glucanotransferase [Thermotoga neopolitana]			
394	10104483 (787, 788)	Novel Protein sim. GBank gl2677790 (U70327) - unknown	Contains protein domain (PF00047) -	UNCLASSIFIED	264693
395	8225070 (789, 790)	(Paretopus polyactis)	Immunoglobulin domain	UNCLASSIFIED	264556
396	24362224 (791, 792)	Novel Protein sim. GBank gl1518468 (U45988) -		struct	264766
397	80417014 (793, 794)	gl45074(pj14A00308.1) p14A00308.1 [Vesikot-Mdrich			265007, 265009, 264608, 264556, 264620,
398	9123517 (795, 796)	gl45074(pj14A00308.1) p14A00308.1 [Vesikot-Mdrich	Contains protein domain (PF00153) -	transport	18103594, 22278095, 22278096, 5694075,
		gl45074(pj14A00308.1) p14A00308.1 [Vesikot-Mdrich	Mitochondrial carrier proteins		22278099, 264259, 26331824, 26331826,
		gl45074(pj14A00308.1) p14A00308.1 [Vesikot-Mdrich			264905, 264906, 265007, 265008, 265009,
		gl45074(pj14A00308.1) p14A00308.1 [Vesikot-Mdrich			21900764, 33657084, 265017, 264448,
		gl45074(pj14A00308.1) p14A00308.1 [Vesikot-Mdrich			264288, 264768, 21906765, 21906766,
		gl45074(pj14A00308.1) p14A00308.1 [Vesikot-Mdrich			21906767, 265020, 265021, 33657023,
		gl45074(pj14A00308.1) p14A00308.1 [Vesikot-Mdrich			33657106, 264628, 36696423, 36696455,
		gl45074(pj14A00308.1) p14A00308.1 [Vesikot-Mdrich			264952, 18108380, 264567, 18108351

399	80053278 (797, 798)	Novel Protein sim. GBank g13356951(p9jBjA31995) - (A3013974) glycerol kinase (Pseudomonas) kinase	Contains protein domain (PF00370) - FCGY family of carbohydrate kinases	kinase	264592, 264595
400	94117480 (798, 800)	Novel Protein sim. GBank g17284355(p9jBj92ALUS_HUMAN - III) ALU SUBFAMILY SC WARNING ENTRY III	Contains protein domain (PF00563) - (citrulline Rich repeat	cadherin	18108394, 56129375, 22278995, 22278997, 22278999, 264259, 26331824, 265006, 265007, 265008, 60432229, 3357402, 265010, 265011, 265012, 265013, 265014, 265015, 265016, 265017, 265018, 265019, 18108351, 18108352, 265020, 265021, 265022, 264691, 264692, 33557023, 18108370, 63274791, 264634, 264636, 60170394, 56129323
401	11397491 (801, 802)	Novel Protein sim. GBank g14924822(p9jA333527.1)AF13211 - (AF132117) FnuA (Staphylococcus aureus)		transport	264594
402	95420294 (803, 804)	Novel Protein sim. GBank g15890437(p9jBjAA83027.1) - (A3026998) KIAA1075 protein [Homo sapiens]	Contains protein domain (PF00017) - Src homology domain 2	phosphatase	65274572, 56182575, 35696286, 22278996, 22278998, 264093, 264259, 26531822, 26531824, 26531825, 26531826, 60432289, 26531827, 26531828, 264908, 264907, 264909, 265006, 264511, 265007, 265008, 264910, 264991, 33657402, 60433356, 60433358, 264596, 21900754, 52844266, 265010, 265011, 87108359, 265017, 265018, 265019, 264681, 18108351, 264682, 264448, 264683, 264684, 264685, 264757, 264686, 264687, 264688, 264689, 264690, 264691, 264692, 264693, 264694, 264695, 264696, 264697, 264698, 264699, 264700, 264701, 264702, 264703, 264704, 264705, 264706, 264707, 264708, 264709, 264710, 264711, 264712, 264713, 264714, 264715, 264716, 264717, 264718, 264719, 264720, 264721, 264722, 264690, 264693, 65274620, 35695763, 264629, 18108370, 264629, 18108379, 35696423, 55811576, 264635, 264636, 264557, 264639, 18108385, 22279002, 264563, 264564, 264565, 264566, 264768, 264632, 264639, 264563
403	80439913 (805, 806)	Novel Protein sim. GBank g1266164(p9jBjCAA15755) - (A1009195) DnaE2 (Mycobacterium tuberculosis)		UNCLASSIFIED	264682
404	8139593 (807, 808)			polymerase	265009, 264682
405	75471200 (805, 810)		Contains protein domain (PF00159) - Pancreatic hormone peptides	UNCLASSIFIED	18108357, 264693
406	79654172 (811, 812)			UNCLASSIFIED	
407	80478229 (813, 814)			UNCLASSIFIED	264769
408	80079556 (815, 816)			UNCLASSIFIED	264600
409	5640527 (817, 818)	Novel Protein sim. GBank g13047117 (A1058919) - similar to ATP-dependent RNA helicases (Arabidopsis thaliana)		helicase	264259

410	95357456 (815, 820)	Novel Protein sim. GBank g14750 (p6b)BA061594 - (Q29801) Unknown [Mus musculus]	UNCLASSIFIED	264489, 52645355, 52646824, 56181696, 5666296, 52645080, 29331322, 29331824, 56182161, 29331825, 60424269, 35699032, 33656970, 264508, 264509, 264905, 264965, 264970, 264985, 32644045, 264908, 264510, 265007, 264512, 265003, 264910, 33657402, 264756, 56463177, 55611386, 265010, 265011, 265017, 264694, 265018, 55611150, 264743, 264745, 264746, 264747, 264748, 264749, 264750, 52644229, 211679, 265020, 265021, 264534, 52644150, 264693, 33657003, 65274620, 33657100, 33657182, 27486261, 35695753, 264624, 3649923, 80411528, 18108376, 263978, 35696423, 35655855, 264632, 264634, 264635, 264637, 264638, 264639, 264639, 56182323, 264558, 60432113, 22279002, 264563, 264565, 264486	UNCLASSIFIED	264769
411	86501570 (821, 822)	Novel Protein sim. GBank g1326176 (emb)CA089971 - (Z95556) htpX [Mycobacterium tuberculosis]	UNCLASSIFIED	264307, 264510, 263973, 222719002		
412	80241602 (823, 824)	Novel Protein sim. GBank	eph	264605		
413	11076446 (825, 826)	Novel Protein sim. GBank	dehydrogenase	18108374, 264760, 264769, 264602, 264633, 264603, 264909, 264805		
414	82050554 (827, 828)	g11280369 (p207)T0001_AZOV1 - 2-OXOGLUTARATE DEHYDROGENASE E1 COMPONENT (ALPHA-KETOGLUTARATE DEHYDROGENASE)	UNCLASSIFIED	264908, 87168518		
415	84433144 (829, 830)	Novel Protein sim. GBank g146883350 (p1)AA031273.1 (AF13202) (Rhopilin [Drosophila melanogaster])	UNCLASSIFIED	264488, 264600, 264602, 267154, 264636		
416	80402775 (831, 832)	Novel Protein sim. GBank g12555172 (AF025433) - ArcC, carbamate kinase [Rizobium etli]	kinase	264605		
417	20153797 (833, 834)	Novel Protein sim. GBank g12555172 (AF025433) - ArcC, carbamate kinase [Rizobium etli]	kinase	264605		
418	94125841 (835, 836)	g11708171 (p1)P531_TMYX2_XANOR - MODIFICATION METHYLASE XORR1 (CITROSINE-SPECIFIC METHYLTRANSFERASE XORR1) (M_XORR1)	UNCLASSIFIED	264605		
419	95314273 (837, 838)	Novel Protein sim. GBank g1326176 (emb)CA061594 - (Q29801) Unknown [Mus musculus]	UNCLASSIFIED	264605		
420	37039549 (839, 840)	Novel Protein sim. GBank g1326176 (emb)CA061594 - (Q29801) Unknown [Mus musculus]	UNCLASSIFIED	264605		
421	95292942 (841, 842)	Novel Protein sim. GBank g12616942 (emb)CA4175801 - (A021959) hypothetical protein Rv2498 [Mycobacterium tuberculosis]	UNCLASSIFIED	264605		
422	79477203 (843, 844)	Novel Protein sim. GBank g12616942 (emb)CA4175801 - (A021959) hypothetical protein Rv2498 [Mycobacterium tuberculosis]	UNCLASSIFIED	264605		
423	75040948 (845, 846)	Novel Protein sim. GBank g12616942 (emb)CA4175801 - (A021959) hypothetical protein Rv2498 [Mycobacterium tuberculosis]	UNCLASSIFIED	264605		

424	78956537 (847, 848)	Novel Protein sim. GBank g14261814[refNP_004977.1]pKTK1 - kinesin (receptor)		stand	265019
425	80431430 (849, 850)	Novel Protein sim. GBank g1703701[bbq178462 - Kinesin-related protein (R373, 850), Peptide Partal, 187 aa]	Contains protein domain (PF00223) - Kinesin motor domain	stand	264909, 265007, 35811396, 264768, 35810764
426	80054322 (851, 852)	Novel Protein sim. GBank g12318293[refP29829]COBN - PSEDE - COBN PROTEIN		UNCLASSIFIED	264605, 264559, 264603, 264536
427	80057232 (853, 854)	Novel Protein sim. GBank g12318293[refP29829]COBN - PSEDE - COBN PROTEIN		UNCLASSIFIED	264603, 264536
428	79407788 (855, 856)	Novel Protein sim. GBank g1812861p[IS22697 - extenain - Vihor carteri (fragment)]		UNCLASSIFIED	264603, 264536
429	80091232 (857, 858)	Novel Protein sim. GBank g1812861p[IS22697 - extenain - Vihor carteri (fragment)]		UNCLASSIFIED	264603, 264536
430	80554192 (859, 860)	Novel Protein sim. GBank g1812861p[IS22697 - extenain - Vihor carteri (fragment)]		UNCLASSIFIED	264603, 264536
431	26824249 (861, 862)	Novel Protein sim. GBank g1323355[emb]CAAL1809] - (A02578) d.034P122 (hypothetical Proteine-rich protein from <i>Neisseria meningitidis</i>)		UNCLASSIFIED	264603, 264536
432	11525372 (863, 864)	Novel Protein sim. GBank g1323355[emb]CAAL1809] - (A02578) d.034P122 (hypothetical Proteine-rich protein from <i>Neisseria meningitidis</i>)		UNCLASSIFIED	264603, 264536
433	81494303 (865, 866)	Novel Protein sim. GBank g1323355[emb]CAAL1809] - (A02578) d.034P122 (hypothetical Proteine-rich protein from <i>Neisseria meningitidis</i>)		UNCLASSIFIED	264603, 264536
434	94363323 (867, 868)	Novel Protein sim. GBank g1323355[emb]CAAL1809] - (A02578) d.034P122 (hypothetical Proteine-rich protein from <i>Neisseria meningitidis</i>)	Contains protein domain (PF00169) - PH domain	UNCLASSIFIED	264603, 264536
435	80502738 (869, 870)	Novel Protein sim. GBank g1323355[emb]CAAL1809] - (A02578) d.034P122 (hypothetical Proteine-rich protein from <i>Neisseria meningitidis</i>)		UNCLASSIFIED	264603, 264536
436	41065523 (871, 872)	Novel Protein sim. GBank g1323355[emb]CAAL1809] - (A02578) d.034P122 (hypothetical Proteine-rich protein from <i>Neisseria meningitidis</i>)		UNCLASSIFIED	264603, 264536
437	11399291 (873, 874)	Novel Protein sim. GBank g1323355[emb]CAAL1809] - (A02578) d.034P122 (hypothetical Proteine-rich protein from <i>Neisseria meningitidis</i>)		UNCLASSIFIED	264603, 264536
438	11773835 (875, 876)	Novel Protein sim. GBank g1323355[emb]CAAL1809] - (A02578) d.034P122 (hypothetical Proteine-rich protein from <i>Neisseria meningitidis</i>)		UNCLASSIFIED	264603, 264536
439	80019465 (877, 878)	Novel Protein sim. GBank g1323355[emb]CAAL1809] - (A02578) d.034P122 (hypothetical Proteine-rich protein from <i>Neisseria meningitidis</i>)		UNCLASSIFIED	264603, 264536
440	79841062 (879, 880)	Novel Protein sim. GBank g1323355[emb]CAAL1809] - (A02578) d.034P122 (hypothetical Proteine-rich protein from <i>Neisseria meningitidis</i>)	Contains protein domain (PF00004) - ATPase, associated cellular activities (AAA)	UNCLASSIFIED	264603, 264536
441	20396935 (881, 882)	Novel Protein sim. GBank g1323355[emb]CAAL1809] - (A02578) d.034P122 (hypothetical Proteine-rich protein from <i>Neisseria meningitidis</i>)		UNCLASSIFIED	264603, 264536
442	85201056 (883, 884)	Novel Protein sim. GBank g1323355[emb]CAAL1809] - (A02578) d.034P122 (hypothetical Proteine-rich protein from <i>Neisseria meningitidis</i>)		UNCLASSIFIED	264603, 264536
443	82456427 (885, 886)	Novel Protein sim. GBank g1323355[emb]CAAL1809] - (A02578) d.034P122 (hypothetical Proteine-rich protein from <i>Neisseria meningitidis</i>)		UNCLASSIFIED	264603, 264536
444	11394667 (887, 888)	Novel Protein sim. GBank g1323355[emb]CAAL1809] - (A02578) d.034P122 (hypothetical Proteine-rich protein from <i>Neisseria meningitidis</i>)		UNCLASSIFIED	264603, 264536

445	76552709 (885, 890)	Novel Protein sim. GBank g1653127[emoj]AB50697.1] - [A243600] WSC4 homologue (Karyosynonymes laevis)		UNCLASSIFIED	264693
446	79810937 (891, 892)	Novel Protein sim. GBank g1653127[emoj]AB50697.1] - [A243600] WSC4 homologue (Karyosynonymes laevis)		UNCLASSIFIED	264509
447	80438888 (893, 894)	Novel Protein sim. GBank g1653127[emoj]AB50697.1] - [A243600] WSC4 homologue (Karyosynonymes laevis)	Contains protein domain (PF00098) - Zinc finger, C2H2 type	transcription factor	264766, 5811576
448	80238110 (895, 896)	Novel Protein sim. GBank g1653127[emoj]AB50697.1] - [A243600] WSC4 homologue (Karyosynonymes laevis)	Contains protein domain (PF00098) - Zinc finger, C2H2 type	dehydrogenase	264509, 264600, 264603, 264605, 264682, 264766, 18103362, 294634, 18103367
449	20466634 (897, 898)	Novel Protein sim. GBank g1653127[emoj]AB50697.1] - [A243600] WSC4 homologue (Karyosynonymes laevis)	Formyl transferase	polymerase	264603, 264559
450	94631210 (899, 900)	Novel Protein sim. GBank g1653127[emoj]AB50697.1] - [A243600] WSC4 homologue (Karyosynonymes laevis)	Contains protein domain (PF00481) - Protein phosphatase 2C	phosphatase	65274572, 22278988, 26331824, 26331826, 264505, 264610, 264592, 52646317, 765017, 21905787, 55811557, 56526486, 22279002
451	214338609 (901, 902)	Novel Protein sim. GBank g1653127[emoj]AB50697.1] - [A243600] WSC4 homologue (Karyosynonymes laevis)		UNCLASSIFIED	264486
452	10267276 (903, 904)	Novel Protein sim. GBank g1653127[emoj]AB50697.1] - [A243600] WSC4 homologue (Karyosynonymes laevis)		UNCLASSIFIED	264692
453	52560096 (905, 906)	Novel Protein sim. GBank g1653127[emoj]AB50697.1] - [A243600] WSC4 homologue (Karyosynonymes laevis)		UNCLASSIFIED	264907, 264600
454	38523922 (907, 908)	Novel Protein sim. GBank g1653127[emoj]AB50697.1] - [A243600] WSC4 homologue (Karyosynonymes laevis)	transferrase		264603
455	13089602 (909, 910)	Novel Protein sim. GBank g1653127[emoj]AB50697.1] - [A243600] WSC4 homologue (Karyosynonymes laevis)		UNCLASSIFIED	264487
456	75663081 (911, 912)	Novel Protein sim. GBank g1653127[emoj]AB50697.1] - [A243600] WSC4 homologue (Karyosynonymes laevis)		UNCLASSIFIED	264691
457	76831273 (913, 914)	Novel Protein sim. GBank g1653127[emoj]AB50697.1] - [A243600] WSC4 homologue (Karyosynonymes laevis)		UNCLASSIFIED	264605
458	79881227 (915, 916)	Novel Protein sim. GBank g1653127[emoj]AB50697.1] - [A243600] WSC4 homologue (Karyosynonymes laevis)	Contains protein domain (PF00359) - PDZ domain (Also known as DHR or GLGF)	kinase	55812038, 265010, 265018, 264681
459	80567359 (917, 918)	Novel Protein sim. GBank g1653127[emoj]AB50697.1] - [A243600] WSC4 homologue (Karyosynonymes laevis)	Contains protein domain (PF00359) - PDZ domain (Also known as DHR or GLGF)	kinase	22278997, 264239, 26331826, 265018, 264446, 264389, 21906795, 36966423
460	76245980 (919, 920)	Novel Protein sim. GBank g1653127[emoj]AB50697.1] - [A243600] WSC4 homologue (Karyosynonymes laevis)	domain C1 domain	UNCLASSIFIED	264606
461	98287618 (921, 922)	Novel Protein sim. GBank g1653127[emoj]AB50697.1] - [A243600] WSC4 homologue (Karyosynonymes laevis)		synthase	264602, 264605, 264768, 264789, 265021, 33031023, 264539

462	7600589 (923, 924)	Novel Protein sim. Gbank g1346019p p4397 PTT1_XANCP- MULTIPHOSPHORYL TRANSFER PROTEIN (MTP) CONTAINS PHOSPHATASE AND PHOSPHATE PROTEIN SYSTEM TRANSFERASE (PHOSPHATASE) SYSTEM ENZYME IN PHOSPHOCARRIER PROTEIN (MTP) (PROTEIN A); PTS SYSTEM, FRUCTOSE-SPECIFIC IIA COMPONENT	Contains protein domain (PF00391) PEP-utilizing enzymes	UNCLASSIFIED	264907
463	70706417 (925, 926)	Novel Protein sim. Gbank g1854045 emb CA583371- (A03413) U98 [Human herpesvirus 6]		UNCLASSIFIED	264905, 264906, 264908, 264909, 264907, 264910, 264991, 264955, 265011, 264632, 284435, 264636, 264637, 264638, 264639
464	82340151 (927, 928)	Novel Protein sim. Gbank g1568577 emb CA852137.1]- (A1224352) calpain [Homo sapiens]	Contains protein domain (PF00648) Calpain family cysteine protease	UNCLASSIFIED	264634
465	83005730 (929, 930)	Novel Protein sim. Gbank g1100617 emb CA036470]- (Z434395) rpsC [Mycobacterium tuberculosis]	Contains protein domain (PF00417) Ribosomal protein S3, N-terminal domain	UNCLASSIFIED	265017, 2190674, 265020
466	20460645 (931, 932)	Novel Protein sim. Gbank g1548705 p 38549 PRBS9_BACSU - D-RIBOSE-BINDING PROTEIN PRECURSOR		UNCLASSIFIED	264605, 264559
467	86409035 (933, 934)	Novel Protein sim. Gbank g1211402 emb CA008957]- (Z95559) grcC [Mycobacterium tuberculosis]		UNCLASSIFIED	264764
468	52562208 (935, 936)	Novel Protein sim. Gbank g12094459 emb CA017347]- (A0211929) cobQ [Mycobacterium tuberculosis]		UNCLASSIFIED	264682
469	18520527 (937, 938)	Novel Protein sim. Gbank g114921 p p17447 BETT_ECOLI - HIGH-AFFINITY CHOLINE TRANSPORT PROTEIN		UNCLASSIFIED	264498
470	80502756 (939, 940)	Novel Protein sim. Gbank g114921 p p17447 BETT_ECOLI - HIGH-AFFINITY CHOLINE TRANSPORT PROTEIN		UNCLASSIFIED	264602, 264789
471	17937351 (941, 942)	Novel Protein sim. Gbank g114921 p p17447 BETT_ECOLI - HIGH-AFFINITY CHOLINE TRANSPORT PROTEIN		UNCLASSIFIED	265019
472	80047458 (943, 944)	Novel Protein sim. Gbank g1862243 (L10509) - Grap1 gene product [Mus musculus]		UNCLASSIFIED	264566, 264685, 264567
473	20538793 (945, 946)	Novel Protein sim. Gbank g1211402 emb CA008957]- (Z95559) grcC [Mycobacterium tuberculosis]		UNCLASSIFIED	264499
474	80593385 (947, 948)	Novel Protein sim. Gbank g12094459 emb CA017347]- (A0211929) cobQ [Mycobacterium tuberculosis]		UNCLASSIFIED	264907, 264908, 264632, 264385
475	82446465 (949, 950)	Novel Protein sim. Gbank g114921 p p17447 BETT_ECOLI - HIGH-AFFINITY CHOLINE TRANSPORT PROTEIN		UNCLASSIFIED	264497, 264908, 264631, 265009, 264762, 264498, 264636, 264639
476	94143857 (951, 952)	Novel Protein sim. Gbank g1543058 p p1006329-1 pGAC1 - glona amplified on chromosome 1 protein (leucine-rich)	Contains protein domain (PF00560) Leucine Rich Repeat	UNCLASSIFIED	65274572, 004330049, 264259, 264503, 5264045, 55812038, 264759, 265011, 264498, 264686, 52644229, 65274791, 264638, 264696
477	70174833 (953, 954)	Novel Protein sim. Gbank g11127551 (U186309) - orf2 [Bitterchocotus kakulensis]		UNCLASSIFIED	264638
478	78633483 (955, 956)	Novel Protein sim. Gbank g11127551 (U186309) - orf2 [Bitterchocotus kakulensis]		UNCLASSIFIED	264686, 264693
479	80189746 (957, 958)	Novel Protein sim. Gbank g11127551 (U186309) - orf2 [Bitterchocotus kakulensis]		UNCLASSIFIED	264886, 36695565, 265008, 264631, 264910, 264632, 264638, 265018, 264389, 264909
480	73650729 (959, 960)	Novel Protein sim. Gbank g11127551 (U186309) - orf2 [Bitterchocotus kakulensis]		UNCLASSIFIED	264389
481	70174833 (961, 962)	Novel Protein sim. Gbank g11127551 (U186309) - orf2 [Bitterchocotus kakulensis]		UNCLASSIFIED	264693
482	83606811 (963, 964)	Novel Protein sim. Gbank g11127551 (U186309) - orf2 [Bitterchocotus kakulensis]		UNCLASSIFIED	264909, 264686, 264788, 264603, 5581178, 56112323, 18109385

483	20253308 (965, 966)	Novel Protein sim. GBank glp2104303(emb)(C1608532) - (Z45387) hypothetical protein Rv2510c [Mycobacterium tuberculosis]	Contains protein domain (PF00534) - Glycosyl transferases group 1	264600	
484	11818348 (967, 968)	Novel Protein sim. GBank glj3450883 (AF003334) - fibron (Austriana penny)		UNCLASSIFIED	264594
485	80191234 (969, 970)	Novel Protein sim. GBank glj54272(emb)(CA944526.1) - (AL078618) nuof. NADH dehydrogenase subunit [Streptomyces coelicolor]		UNCLASSIFIED	264604
486	80059042 (971, 972)	Novel Protein sim. GBank glj54272(emb)(CA944526.1) - (AL078618) nuof. NADH dehydrogenase subunit [Streptomyces coelicolor]		UNCLASSIFIED	264604
487	11813339 (973, 974)	Novel Protein sim. GBank glj54272(emb)(CA944526.1) - (AL078618) nuof. NADH dehydrogenase subunit [Streptomyces coelicolor]		UNCLASSIFIED	264604
488	91222383 (975, 976)	Novel Protein sim. GBank glj54272(emb)(CA944526.1) - (AL078618) nuof. NADH dehydrogenase subunit [Streptomyces coelicolor]		UNCLASSIFIED	264604
489	10657710 (977, 978)	Novel Protein sim. GBank glj54272(emb)(CA944526.1) - (AL078618) nuof. NADH dehydrogenase subunit [Streptomyces coelicolor]		UNCLASSIFIED	264604
490	95361124 (979, 980)	Novel Protein sim. GBank glj54272(emb)(CA944526.1) - (AL078618) nuof. NADH dehydrogenase subunit [Streptomyces coelicolor]		UNCLASSIFIED	264604
491	80458412 (981, 982)	Novel Protein sim. GBank glj54272(emb)(CA944526.1) - (AL078618) nuof. NADH dehydrogenase subunit [Streptomyces coelicolor]		UNCLASSIFIED	264604
492	97421264 (983, 984)	Novel Protein sim. GBank glj54272(emb)(CA944526.1) - (AL078618) nuof. NADH dehydrogenase subunit [Streptomyces coelicolor]		UNCLASSIFIED	264604
493	11692942 (985, 986)	Novel Protein sim. GBank glj54272(emb)(CA944526.1) - (AL078618) nuof. NADH dehydrogenase subunit [Streptomyces coelicolor]		UNCLASSIFIED	264604
494	87726604 (987, 988)	Novel Protein sim. GBank glj54272(emb)(CA944526.1) - (AL078618) nuof. NADH dehydrogenase subunit [Streptomyces coelicolor]		UNCLASSIFIED	264604
495	80235599 (989, 990)	Novel Protein sim. GBank glj54272(emb)(CA944526.1) - (AL078618) nuof. NADH dehydrogenase subunit [Streptomyces coelicolor]		UNCLASSIFIED	264604
496	78985624 (991, 992)	Novel Protein sim. GBank glj54272(emb)(CA944526.1) - (AL078618) nuof. NADH dehydrogenase subunit [Streptomyces coelicolor]		UNCLASSIFIED	264604
497	78945621 (993, 994)	Novel Protein sim. GBank glj54272(emb)(CA944526.1) - (AL078618) nuof. NADH dehydrogenase subunit [Streptomyces coelicolor]		UNCLASSIFIED	264604

517	9522294 (1033, 1034)	Novel Protein sim. GBank g12933605 (AE000725) - ribose 5 phosphate isomerase B [Aquifex aeolicus]	Isomerase	265018, 264605, 264764, 264766, 264687, 264691, 264585
518	8491831 (1035, 1036)	Novel Protein sim. GBank g1854055 (emj) CAAG33371 - (283413) U88 [Human Heparanase 8]	UNCLASSIFIED	284487
519	91677888 (1037, 1038)	Novel Protein sim. GBank g19589060 (g19589073) - (46024079) 6120 [Homo sapiens]	dna_mg_bind	52644507, 22278897, 22278988, 60432048, 264239, 52645090, 29331824, 66714117, 6042288, 29331828, 35995002, 264905, 29331830, 65712502, 264911, 265007, 264991, 60432228, 33657402, 60433438, 264993, 35995006, 264994, 87169474, 87169459, 3585017, 265018, 26496, 264919, 264681, 264448, 264589, 264288, 264685, 264687, 21906765, 21906766, 21906757, 21906768, 21906761, 60170615, 35657033, 264692, 52645129, 35657109, 27486262, 27486264, 35693763, 81103370, 264629, 52644332, 56183323, 264638, 83373044, 18108385, 56526486, 50432113
520	78869188 (1039, 1040)	Novel Protein sim. GBank g11091209 (p1463) CTPA_MYCLE - CATION-PROTON-ATPASE A	transport	264769
521	11076827 (1041, 1042)	Novel Protein sim. GBank g11172869 (p1433) IBBSK_H4SEN_RIBOXINASE	kinase	264605, 264768
522	80435060 (1043, 1044)	Novel Protein sim. GBank g1213243 (p1) S51028 - hypothetical protein YPI_238c - yeast [Saccharomyces cerevisiae]	UNCLASSIFIED	264629
523	18355013 (1045, 1046)	Novel Protein sim. GBank g1213243 (p1) S51028 - hypothetical protein YPI_238c - yeast [Saccharomyces cerevisiae]	ATPase associated	264032, 264598, 265011
524	80261805 (1047, 1048)	Novel Protein sim. GBank g10336009 (p1) BAA35138 - (AB012308) B2HC [Anthracinus crassipalpis]	UNCLASSIFIED	264807
525	79510046 (1049, 1050)	Novel Protein sim. GBank g14106810 (emb) CAA21363 - (AL031856) ORF42, lens-386 aa, similarity to an aminotransferase, in P95957 Sulfolobus solfataricus, (401 aa), 33.1% identity in 393 aa overlap, Fasta scores: opt-468, E1): 8.5e-24, in G64602 R. norvegicus, (425 aa), 28.6% ident...	UNCLASSIFIED	264758
526	38827630 (1051, 1052)	Novel Protein sim. GBank g1731046 (p1) X41614, ECOLI - MANNONATE DEHYDRATASE [D-MANNONATE HYDROLYASE] (X00513) NuaA protein (nuaA) [Escherichia coli]	UNCLASSIFIED	264769
527	80504739 (1053, 1054)	Novel Protein sim. GBank g1731046 (p1) X41614, ECOLI - MANNONATE DEHYDRATASE [D-MANNONATE HYDROLYASE] (X00513) NuaA protein (nuaA) [Escherichia coli]	UNCLASSIFIED	55182575, 265017, 265018
528	80544113 (1055, 1056)	Novel Protein sim. GBank g1731046 (p1) X41614, ECOLI - MANNONATE DEHYDRATASE [D-MANNONATE HYDROLYASE] (X00513) NuaA protein (nuaA) [Escherichia coli]	hydrolase	265019
529	17936810 (1057, 1058)	Novel Protein sim. GBank g1731046 (p1) X41614, ECOLI - MANNONATE DEHYDRATASE [D-MANNONATE HYDROLYASE] (X00513) NuaA protein (nuaA) [Escherichia coli]	UNCLASSIFIED	264687
530	10887336 (1059, 1060)	Novel Protein sim. GBank g15257640 (emb) CAB49758.1 - (AL080170) hypothetical protein [Homo sapiens]	UNCLASSIFIED	264555, 264556, 264557, 264558, 18108385
531	80226576 (1061, 1062)	Novel Protein sim. GBank g15257640 (emb) CAB49758.1 - (AL080170) hypothetical protein [Homo sapiens]	UNCLASSIFIED	264498, 264499, 264499, 264502, 264766, 265021, 264990, 263978, 264558
532	50933444 (1063, 1064)	Novel Protein sim. GBank g15257640 (emb) CAB49758.1 - (AL080170) hypothetical protein [Homo sapiens]	UNCLASSIFIED	

533	87761531 (1065, 1066)	Novel Protein sm. GBank gl4663636/pJAD31953.1(AF11229) - (AF112299) integral inner nuclear membrane protein MANT [Homo sapiens]		UNCLASSIFIED	264907, 264909, 264786, 35695917, 264630, 264555
534	82362624 (1067, 1068)	Novel Protein sm. GBank gl2495352/pmcj040406.1) - (AF112299) integral inner nuclear membrane protein MANT [Homo sapiens]		UNCLASSIFIED	264925, 264911, 264601, 264602, 264605, 264782, 264786, 265020, 264693, 264636, 264606
535	75841650 (1069, 1070)	Novel Protein sm. GBank gl2495352/pmcj040406.1) - (AF112299) integral inner nuclear membrane protein MANT [Homo sapiens] (Z69723) similar to cAMP-dependent protein kinase cDNA EST EMBL:700719 comes from this gene cDNA EST y46548.3 comes from this gene cDNA EST y46548.5 comes from this gene cDNA EST y44324.3 comes from this gene cDNA EST y...	Cyclins protein domain (PF00063) Eukaryotic protein kinase domain	ATPase-associated	
536	76907207 (1071, 1072)	Novel Protein sm. GBank gl2495628/pJ5757YOH1_SERUA - HYPOTHETICAL 10.1 KD PROTEIN IN BIOA 57EGION		reductase	16108376, 264605, 264606, 264607, 264609
537	94147448 (1073, 1074)	Novel Protein sm. GBank gl1349201/pJ1987SSGP_VOLCA - SULFATED SURFACE GLYCOPROTEIN 185 (SSG 185)	Contains protein domain (PF00599) - PDZ domain (Also known as DMR or GLGF)	collagen	265006, 264605, 65274781
538	87821963 (1075, 1076)	Novel Protein sm. GBank gl3462214/dbjBA334480.1) - (A3016303) KIAA0760 protein [Homo sapiens]			29331822, 29331824, 29331825, 29331826, 29331827, 264606, 5964045, 33657402, 265017, 264782, 264683, 264286, 264685, 21006765, 39656763, 264558, 60170394, 264559, 22278002
539	26306269 (1077, 1078)	Novel Protein sm. GBank gl2495352/pmcj040406.1) - (AF112299) integral inner nuclear membrane protein MANT [Homo sapiens] ACETYLTROANSFERASE		histone	264602, 265019
540	76637077 (1079, 1080)	Novel Protein sm. GBank gl3462214/dbjBA334480.1) - (A3016303) KIAA0760 protein [Homo sapiens]		transcriptfactor	264683
541	87762268 (1081, 1082)	Novel Protein sm. GBank gl5042272/jmbjC844526.1) - (A1078619) nuf. NADH dehydrogenase subunit [Streptomyces coelicolor]	Contains protein domain (PF00069) - Zinc finger, C2H2 type		18108394, 22276967, 22276968, 264259, 264112, 265008, 33657402, 55817038, 52546317, 265017, 21006765, 264693, 55811576, 264535, 56526486, 264566
542	86226636 (1083, 1084)	Novel Protein sm. GBank gl5042272/jmbjC844526.1) - (A1078619) nuf. NADH dehydrogenase subunit [Streptomyces coelicolor]		dehydrogenase	264910, 265018, 264689, 264638, 264486
543	76796290 (1085, 1086)	Novel Protein sm. GBank gl2751380/jmbjCA115894) - (A1021164) hypothetical protein Ry164 [Mycobacterium tuberculosis]		UNCLASSIFIED	264602, 264908
544	20437191 (1087, 1088)	Novel Protein sm. GBank gl2751380/jmbjCA115894) - (A1021164) hypothetical protein Ry164 [Mycobacterium tuberculosis]		UNCLASSIFIED	264605
545	86434504 (1089, 1090)	Novel Protein sm. GBank gl4682214/pJAD33237.1(JAF14744) - (AF147449) penicillin binding protein 1B (penicillinamonds acetylase)			264786, 264634, 264607, 264592, 264609
546	80249016 (1091, 1092)	Novel Protein sm. GBank gl1350655/pJ1987SSGP_VOLCA - SULFATED SURFACE GLYCOPROTEIN 185 (SSG 185)			264600, 264602, 21006765
547	11077553 (1093, 1094)	Novel Protein sm. GBank gl1350655/pJ1987SSGP_VOLCA - SULFATED SURFACE GLYCOPROTEIN 185 (SSG 185)		mapolymerase	264604
548	82114926 (1095, 1096)	Novel Protein sm. GBank gl2330021 (AF1019250) - kinesin-related protein; KRSP, Costal2 [Drosophila melanogaster]		UNCLASSIFIED	264486, 264905, 264910, 264780, 264693, 264639, 264593, 264684

549	95421904 (1097, 1098)	Novel Protein sim. GBank gH433746(gH4AD18133) - (AF056195) neuroblastoma-amplified protein [Homo sapiens]	UNCLASSIFIED	264488, 65274372, 18103368, 22278995, 22278996, 22278997, 22278998, 22278999, 264259, 29331824, 60741117, 29331825, 29331826, 35696052, 265007, 265008, 264910, 264592, 33557402, 33109954, 265017, 265018, 265019, 18108351, 264448, 264754, 264368, 264280, 254766, 264586, 264658, 21906765, 21906766, 21906767, 21906768, 21906769, 265020, 264691, 33557023, 264692, 264693, 65274620, 3245129, 33637106, 27482261, 27482682, 27482264, 33637349, 55811576, 18108387, 264693, 22278992, 264693
550	10885616 (1099, 1100)	Novel Protein sim. GBank gI322883(gI3P94989)(SYFB, MYCTU - PHENYLALANYL-TRNA SYNTHETASE BETA CHAIN (PHENYLALANINE--TRNA LIGASE BETA CHAIN) (PIERS)	UNCLASSIFIED	264908, 264909, 264768
551	80439980 (1101, 1102)	Novel Protein sim. GBank gI322883(gI3P94989)(SYFB, MYCTU - PHENYLALANYL-TRNA SYNTHETASE BETA CHAIN (PHENYLALANINE--TRNA LIGASE BETA CHAIN) (PIERS)	UNCLASSIFIED	264689, 264639, 264563
552	94872870 (1103, 1104)	Novel Protein sim. GBank gI322883(gI3P94989)(SYFB, MYCTU - PHENYLALANYL-TRNA SYNTHETASE BETA CHAIN (PHENYLALANINE--TRNA LIGASE BETA CHAIN) (PIERS)	UNCLASSIFIED	55811957, 264628
553	80106002 (1105, 1106)	Novel Protein sim. GBank gI322883(gI3P94989)(SYFB, MYCTU - PHENYLALANYL-TRNA SYNTHETASE BETA CHAIN (PHENYLALANINE--TRNA LIGASE BETA CHAIN) (PIERS)	UNCLASSIFIED	264696
554	79613379 (1107, 1108)	Novel Protein sim. GBank gI322883(gI3P94989)(SYFB, MYCTU - PHENYLALANYL-TRNA SYNTHETASE BETA CHAIN (PHENYLALANINE--TRNA LIGASE BETA CHAIN) (PIERS)	UNCLASSIFIED	264762
555	70998347 (1109, 1110)	Novel Protein sim. GBank gI322883(gI3P94989)(SYFB, MYCTU - PHENYLALANYL-TRNA SYNTHETASE BETA CHAIN (PHENYLALANINE--TRNA LIGASE BETA CHAIN) (PIERS)	UNCLASSIFIED	264508, 264505, 264559
556	20457127 (1111, 1112)	Novel Protein sim. GBank gI322883(gI3P94989)(SYFB, MYCTU - PHENYLALANYL-TRNA SYNTHETASE BETA CHAIN (PHENYLALANINE--TRNA LIGASE BETA CHAIN) (PIERS)	UNCLASSIFIED	264488
557	19523405 (1113, 1114)	Novel Protein sim. GBank gI322883(gI3P94989)(SYFB, MYCTU - PHENYLALANYL-TRNA SYNTHETASE BETA CHAIN (PHENYLALANINE--TRNA LIGASE BETA CHAIN) (PIERS)	UNCLASSIFIED	264602
558	20724428 (1115, 1116)	Novel Protein sim. GBank gI322883(gI3P94989)(SYFB, MYCTU - PHENYLALANYL-TRNA SYNTHETASE BETA CHAIN (PHENYLALANINE--TRNA LIGASE BETA CHAIN) (PIERS)	UNCLASSIFIED	264634

360	30066533 (1118, 1120)	Novel Protein sim. GBank glt48295jpl03193Y41R, RHISN - PROBABLE PEPTIDE ABC TRANSPORTER ATP-BINDING PROTEIN Y41R	Contains protein domain (PF00009) - transport ABC transporter	18103396, 264906, 264602, 264604, 18103374
361	20292187 (1121, 1122)	Novel Protein sim. GBank gtl104955 (A1042776) - poly(hydroxyacetic acid) granule associated protein Gx2	UNCLASSIFIED	264600
362	11699161 (1123, 1124)	Novel Protein sim. GBank gtl104955 (A1042776) - poly(hydroxyacetic acid) granule associated protein Gx2	UNCLASSIFIED	264600
363	79761420 (1125, 1126)	Novel Protein sim. GBank gtl104955 (A1042776) - poly(hydroxyacetic acid) granule associated protein Gx2	UNCLASSIFIED	264910, 264691
364	55715360 (1127, 1128)	Novel Protein sim. GBank gtl104955 (A1042776) - poly(hydroxyacetic acid) granule associated protein Gx2	dehydrogenase	264592
365	50485818 (1129, 1130)	Novel Protein sim. GBank gtl104955 (A1042776) - poly(hydroxyacetic acid) granule associated protein Gx2	synthase	265010
366	94323888 (1131, 1132)	Novel Protein sim. GBank gtl104955 (A1042776) - poly(hydroxyacetic acid) granule associated protein Gx2	helicase	264900, 264510, 265000, 264910, 264758, 264600, 264602, 264604, 264605, 264768, 264637, 264698, 39655917, 264693, 85274620, 264488
367	75506955 (1133, 1134)	Novel Protein sim. GBank gtl104955 (A1042776) - poly(hydroxyacetic acid) granule associated protein Gx2	UNCLASSIFIED	264681, 264691, 264593
368	94681793 (1135, 1136)	Novel Protein sim. GBank gtl104955 (A1042776) - poly(hydroxyacetic acid) granule associated protein Gx2	dehydrogenase	264699
369	35506897 (1137, 1138)	Novel Protein sim. GBank gtl104955 (A1042776) - poly(hydroxyacetic acid) granule associated protein Gx2	Malic enzyme	264565
370	76375627 (1139, 1140)	Novel Protein sim. GBank gtl104955 (A1042776) - poly(hydroxyacetic acid) granule associated protein Gx2	Contains protein domain (PF00319) - ribosomal prot	18108376, 18103397, 264565
371	79793961 (1141, 1142)	Novel Protein sim. GBank gtl104955 (A1042776) - poly(hydroxyacetic acid) granule associated protein Gx2	transport	264907, 264909
372	36956838 (1143, 1144)	Novel Protein sim. GBank gtl104955 (A1042776) - poly(hydroxyacetic acid) granule associated protein Gx2	UNCLASSIFIED	264762
373	20715521 (1145, 1146)	Novel Protein sim. GBank gtl104955 (A1042776) - poly(hydroxyacetic acid) granule associated protein Gx2	UNCLASSIFIED	265007, 264601
374	13521562 (1147, 1148)	Novel Protein sim. GBank gtl104955 (A1042776) - poly(hydroxyacetic acid) granule associated protein Gx2	polymerase	264636
375	13076416 (1149, 1150)	Novel Protein sim. GBank gtl104955 (A1042776) - poly(hydroxyacetic acid) granule associated protein Gx2	polymerase	264697
376	20482246 (1151, 1152)	Novel Protein sim. GBank gtl104955 (A1042776) - poly(hydroxyacetic acid) granule associated protein Gx2	264605	
377	8627102 (1153, 1154)	Novel Protein sim. GBank gtl104955 (A1042776) - poly(hydroxyacetic acid) granule associated protein Gx2	dehydrogenase	35696052, 264636
378	11804437 (1155, 1156)	Novel Protein sim. GBank gtl104955 (A1042776) - poly(hydroxyacetic acid) granule associated protein Gx2	transport	264638
379	11794723 (1157, 1158)	Novel Protein sim. GBank gtl104955 (A1042776) - poly(hydroxyacetic acid) granule associated protein Gx2	transport	264682, 264558

580	80059417 (1159, 1160)				22278960, 35656052, 264555, 264556, 264558
581	79230833 (1161, 1162)	Novel Protein sim. GBank g13433131 (AF043777) - lin	Contains protein domain (PF00047) -	UNCLASSIFIED	
582	80048617 (1163, 1164)	[Oscopista medinagarsis]	Immunoglobulin domain	struct	265008, 264564
583	79321392 (1165, 1166)	g15011820g17728YAIR, ECOLI - HYPOPHETICAL 49.0 KD PROTEIN IN ABPA-CYTOE INTERGENIC REGION		transport	265021, 264555, 264557
584	79845024 (1167, 1168)	Novel Protein sim. GBank g1382221dbjBA34470.11- (AB018283) KIAA0750 protein [Homo sapiens]			264584
585	79581454 (1169, 1170)	Novel Protein sim. GBank g14467350jembIAG37574 - (AL035569) probable Glu-RNA Gln amidotransferase subunit [Streptomyces coelicolor]		UNCLASSIFIED	264908, 265007
586	38277486 (1171, 1172)	Novel Protein sim. GBank g15689219dbjBA363043.11- (AB023014) KIAA1091 protein [Homo sapiens]		UNCLASSIFIED	264557
587	80497359 (1173, 1174)			UNCLASSIFIED	264600, 384602, 264605, 264759, 264690,
588	79557239 (1175, 1176)			UNCLASSIFIED	264592
589	79805528 (1177, 1178)			UNCLASSIFIED	22278962, 264807, 264909, 264510, 265009,
590	79815629 (1179, 1180)	Novel Protein sim. GBank g2143393jembIAG093901 - (Z54572) [pos] Mycobacterium tuberculosis		UNCLASSIFIED	265010, 264887, 264789, 35689917,
591	10315340 (1181, 1182)	Novel Protein sim. GBank g451183jgbjA4-D21543.11 - (AF088896) electrotransfer ribiquinone oxidoreductase (Zymomonas mobilis)		UNCLASSIFIED	18108376, 264634, 264636, 264638
592	13889767 (1183, 1184)	Novel Protein sim. GBank g1314901spjP20966jP1TB, ECOLI - FTS SYSTEM, (FRUCTOSE-SPECIFIC IBC COMPONENT (EIIBC-FRU))		UNCLASSIFIED	264906, 264609
593	82346659 (1185, 1186)	Novel Protein sim. GBank g1272368 (L51896) - Lge (Vibrio parahaemolyticus)		UNCLASSIFIED	264691
594	20212392 (1187, 1188)	Novel Protein sim. GBank g1272368 (L51896) - Lge (Vibrio parahaemolyticus)		MHC	263972
595	10164064 (1189, 1190)	Novel Protein sim. GBank g1272368 (L51896) - Lge (Vibrio parahaemolyticus)		dehydrogenase	264511, 264762, 264768, 264486
596	13085170 (1191, 1192)	Novel Protein sim. GBank g1272368 (L51896) - Lge (Vibrio parahaemolyticus)		UNCLASSIFIED	264605
597	80283003 (1193, 1194)	Novel Protein sim. GBank g1272368 (L51896) - Lge (Vibrio parahaemolyticus)		UNCLASSIFIED	264769
598	94140216 (1195, 1196)	Novel Protein sim. GBank g1272368 (L51896) - Lge (Vibrio parahaemolyticus)		UNCLASSIFIED	
599	20385137 (1197, 1198)	Novel Protein sim. GBank g1272368 (L51896) - Lge (Vibrio parahaemolyticus)		UNCLASSIFIED	264758, 55810764, 264585, 264586, 264637,
600	10357663 (1199, 1200)	Novel Protein sim. GBank g1272368 (L51896) - Lge (Vibrio parahaemolyticus)		UNCLASSIFIED	83373044
601	79810404 (1201, 1202)	Novel Protein sim. GBank g1272368 (L51896) - Lge (Vibrio parahaemolyticus)		UNCLASSIFIED	264603
602	10357663 (1199, 1200)	Novel Protein sim. GBank g1272368 (L51896) - Lge (Vibrio parahaemolyticus)		UNCLASSIFIED	264906
603	79810404 (1201, 1202)	Novel Protein sim. GBank g1272368 (L51896) - Lge (Vibrio parahaemolyticus)		UNCLASSIFIED	264510

602	78256602 (1203, 1204)	Novel Protein sim. GBank gll552286 [lpp(AAC34243.1) - AC004411] putative pto kinase [Arctobius thalianus]	Contains protein domain (PF00069) - Eukaryotic protein kinase domain	265007	
603	11466061 (1205, 1206)	Novel Protein sim. GBank	UNCLASSIFIED	264955	
604	8157420 (1207, 1208)	Novel Protein sim. GBank	UNCLASSIFIED	264758	
605	20438657 (1209, 1210)	gll1175322 [lpp(A48177683. HAEIN - HYPOTHETICAL PROTEIN H0883]	UNCLASSIFIED	264605	
606	80334582 (1211, 1212)	Novel Protein sim. GBank	UNCLASSIFIED	264764	
607	95361506 (1213, 1214)	gll02026 [lpp(A493043. IAF15138 - (AF151383) Cdc42 GTPase-activating protein (Ras subfamily)]	UNCLASSIFIED	264508, 264608, 83565542, 264682, 264687, 264688, 264514, 18108376, 35996423, 264638, 264555, 264638	
608	11819888 (1215, 1216)	Novel Protein sim. GBank	UNCLASSIFIED	264682	
609	80894775 (1217, 1218)	Novel Protein sim. GBank	UNCLASSIFIED	264605	
610	79623413 (1219, 1220)	gll248670 [lpp(P55520YALL_RHIN - HYPOTHETICAL 91.8 KD PROTEIN YALL	Contains protein domain (PF00069) - PAS domain	264692	
611	87586206 (1221, 1222)	Novel Protein sim. GBank	UNCLASSIFIED	264692	
612	95287851 (1223, 1224)	Novel Protein sim. GBank gll187356 [lpp(CAB07118) - (Z92772) recD (Mycobacterium tuberculosis)]	Contains protein domain (PF01443) - Viral (Superfamily 1) RNA helicase	264508, 264605, 264907, 264908, 264909, 264511, 264510, 264758, 264604, 264684, 264768, 264689, 264682, 264683, 264638, 264637, 264555	
613	79234715 (1225, 1226)	Novel Protein sim. GBank	nuclease	264600, 264601, 264604, 264769, 264558, 264565	
614	79969348 (1227, 1228)	Novel Protein sim. GBank	UNCLASSIFIED	264565	
615	38586696 (1229, 1230)	Novel Protein sim. GBank gll133995 [lpp(BAA12741) - (D65240) large subunit of NADP-dependent glutamate decarboxylase (E. coli)]	kinase	18108372, 264583	
616	20465331 (1231, 1232)	Novel Protein sim. GBank	synthase	264600, 264602, 264629	
617	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
618	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
619	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
620	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
621	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
622	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
623	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
624	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
625	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
626	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
627	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
628	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
629	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
630	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
631	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
632	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
633	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
634	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
635	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
636	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
637	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
638	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
639	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
640	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
641	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
642	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
643	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
644	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
645	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
646	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
647	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
648	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
649	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
650	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
651	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
652	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
653	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
654	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
655	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
656	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
657	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
658	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
659	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
660	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
661	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
662	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
663	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
664	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
665	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
666	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
667	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
668	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
669	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
670	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
671	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
672	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
673	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
674	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
675	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
676	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
677	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
678	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
679	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
680	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
681	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
682	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
683	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
684	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
685	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
686	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
687	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
688	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
689	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
690	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
691	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
692	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
693	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
694	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
695	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
696	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
697	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
698	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
699	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	
700	91227222 (1233, 1234)	Novel Protein sim. GBank	isomerase	264605	

637	92455196 (1273, 1274)	Novel Protein sim. GBank gll224873jenn[CAB10653] - (Z8262) nckA [Mycobacterium tuberculosis]		nuclease	264906, 264937, 264510, 264511, 264601, 264602, 264603, 264604, 264605, 18108351, 264762, 264766, 264607, 264769, 264608, 264609, 264610, 264611, 264612, 264613, 264559, 18108355
638	14397457 (1275, 1276)	Novel Protein sim. GBank gll467862jenn[CAB1074, 1] - (ALD46845) putative large ATP-binding protein [Streptomyces coelicolor]		stnd	264112, 263974
639	80204210 (1277, 1278)	Novel Protein sim. GBank gll458492jenn[BA76836, 1] - (AB023209) KIAA0992 protein (Homo sapiens)			
640	17929579 (1279, 1280)	Novel Protein sim. GBank gll1432083 (U00961) - homolog to Slp1p, an evolutionarily conserved kinetochore protein in budding yeast [Arabidopsis thaliana]	Contains protein domain (PF01466) Slp1 family	mapolymerase	265000, 265010
641	79635398 (1281, 1282)			UNCLASSIFIED	264693
642	15892737 (1283, 1284)			UNCLASSIFIED	264565
643	81516220 (1285, 1286)			UNCLASSIFIED	264906, 264908, 264758, 264288, 264632, 264635, 264639, 264594
644	11751967 (1287, 1288)			UNCLASSIFIED	264684
645	95010907 (1289, 1290)			UNCLASSIFIED	264632, 264162, 264693, 264639, 264559
646	80656063 (1291, 1292)			UNCLASSIFIED	264593, 264586
647	80257085 (1293, 1294)	Novel Protein sim. GBank gll507613jenn[IP_003738, 1] pTINKS - TANKYRASE	Contains protein domain (PF00023) Ank repeat	transcription factor	264906, 264581
648	8007428 (1295, 1296)	Novel Protein sim. GBank gll1044983jenn[169646 - prolamine [Monodonella turbinata, gonads, Peptide, 106 aa]		UNCLASSIFIED	264600
649	80247447 (1297, 1298)			UNCLASSIFIED	263976
650	117768316 (1299, 1300)			UNCLASSIFIED	264686
651	11776932 (1301, 1302)	Novel Protein sim. GBank gll1346916jenn[12283]PURA_ECOLI - ADENYLOSUCCINATE SYNTHETASE (IMP--ASPARTATE LIGASE)		UNCLASSIFIED	264602, 264638
652	85519794 (1303, 1304)	Novel Protein sim. GBank gll1228776jenn[106881]GOC_MYCTU - HYPOTHETICAL		UNCLASSIFIED	264905, 264907, 264909, 263978, 264637
653	82124947 (1305, 1306)	gll1228776jenn[106881]GOC_MYCTU - HYPOTHETICAL		UNCLASSIFIED	22278996, 264510, 264511, 264512, 264993, 21060754, 264603, 264760, 18108316, 264536
654	95010589 (1307, 1308)	Novel Protein sim. GBank gll1303276jenn[26647]PLSC_ECOLI - 1-ACYL-SN-GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE (1-ACP ACYLTRANSFERASE) (1-ACPAT) (LYSOPHOSPHATIDIC ACID ACYLTRANSFERASE) (LPAAT)	Contains protein domain (PF01553) Acyltransferase	transferase	264593, 264595, 264632, 264592
655	79320662 (1309, 1310)			UNCLASSIFIED	
656	80416738 (1311, 1312)			UNCLASSIFIED	264602, 264605, 264766, 264681
657	20611010 (1313, 1314)			UNCLASSIFIED	264557, 264558

556	87761615 (1315, 1316)	Novel Protein sim. GBank gl5668949jpbj(BAA0300.1) - (AB029001) KIAA1078 protein (Homo sapiens)	UNCLASSIFIED	22278596, 60432049, 29331822, 29331824, 29331823, 265007, 265009, 33557402, 3357094, 265017, 264448, 21900768, 263987, 20281149, 18108370, 18108374, 264482
559	87718663 (1317, 1318)	Novel Protein sim. GBank gl218172gipj(48724 - zinc finger protein PZF - mouse)	transcription factor	22278398, 60432049, 60714117, 29331827, 264482, 264476, 36181862, 18108359, 18108365, 18108370, 18108381
560	81807822 (1319, 1320)	Novel Protein sim. GBank gl34180ipj(15401)SACY - BACSU - LEVANSUCRASE AND SUCRASE SYNTHESIS OPERON ANTI-TERMINATOR	UNCLASSIFIED	284753
561	80026023 (1321, 1322)	Novel Protein sim. GBank gl160284qj(16417166) -	UNCLASSIFIED	264505, 265009, 264600, 264602, 264603, 264604, 264605, 32833886, 18108376, 264636, 18108387, 22279000
562	20463731 (1323, 1324)	Novel Protein sim. GBank gl545229gjb(A022450.1)AF11618 - (AF11618) SecA homolog (Actinobacillus sedentis)comitans	UNCLASSIFIED	264605
563	20628060 (1325, 1326)	Novel Protein sim. GBank gl5689250j(16432881.1) - (A024335) similar to orf3 (Cnemonas testaceolona)	dehydrogenase	264605
564	80508512 (1327, 1328)	Novel Protein sim. GBank gl160284qj(16417166) - (A024335) similar to orf3 (Cnemonas testaceolona)	UNCLASSIFIED	264769
565	80279053 (1329, 1330)	Novel Protein sim. GBank gl160284qj(16417166) - (A024335) similar to orf3 (Cnemonas testaceolona)	UNCLASSIFIED	264600
566	80279053 (1329, 1330)	Novel Protein sim. GBank gl160284qj(16417166) - (A024335) similar to orf3 (Cnemonas testaceolona)	isomerase	264600
567	94631802 (1333, 1334)	Novel Protein sim. GBank gl50589351j(16432881.1) - (A024335) similar to orf3 (Cnemonas testaceolona)	glycoprotein	264607, 265007
568	76603142 (1331, 1332)	Novel Protein sim. GBank gl326182j(16432881.1) - (A024335) similar to orf3 (Cnemonas testaceolona)	UNCLASSIFIED	264688, 264602, 264593
569	12967154 (1337, 1338)	Novel Protein sim. GBank gl50589351j(16432881.1) - (A024335) similar to orf3 (Cnemonas testaceolona)	UNCLASSIFIED	264603, 264605, 264760, 264689, 264636, 264637
570	80235549 (1339, 1340)	Novel Protein sim. GBank gl326182j(16432881.1) - (A024335) similar to orf3 (Cnemonas testaceolona)	UNCLASSIFIED	264603, 264605, 264760, 264689, 264636, 264637
571	79601358 (1341, 1342)	Novel Protein sim. GBank gl50589351j(16432881.1) - (A024335) similar to orf3 (Cnemonas testaceolona)	UNCLASSIFIED	264603, 264605, 264760, 264689, 264636, 264637
572	79834371 (1343, 1344)	Novel Protein sim. GBank gl2114430j(527103) - Olf-1 (IEBF)	transcription factor	264759
573	82285768 (1345, 1346)	Novel Protein sim. GBank gl5689250j(16432881.1)AF11618 - (AF11618) SecA homolog (Actinobacillus sedentis)comitans	transcription factor	264759
574	76199259 (1347, 1348)	Novel Protein sim. GBank gl5689250j(16432881.1)AF11618 - (AF11618) SecA homolog (Actinobacillus sedentis)comitans	UNCLASSIFIED	264629

675	87895870 (1349, 1350)	Novel Protein sim. GBank g1498075bpAA035347.1(AE001768) D- asparagine-L-alanine ligase (Thermotoga maritima)	Contains protein domain (PF01820) - D-ala D-ala ligase	UNCLASSIFIED	264488, 22278999, 69714117, 264508, 264511, 265006, 60433438, 264600, 264601, 264602, 264603, 264604, 264605, 264762, 264607, 264768, 60431602, 16106374, 264609, 264630 265010
676	76895607 (1351, 1352)	Novel Protein sim. GBank g1723566bpQ10478YDFT. SCPO - PUTATIVE GLUCOSYLTRANSFERASE C17C9.07		ATPase associated	264591, 264632
677	216414312 (1353, 1354)	Novel Protein sim. GBank g1687208 (Q03376) - dynein heavy chain isoType 5C (Tropaeolus gratia)			264758, 264682, 264557
678	84222200 (1355, 1356)	Novel Protein sim. GBank g11586574 (p172033854) - laminin alpha5 (Mus musculus)	Contains protein domain (PF00053) - laminin Laminin EGF-like (Domains III and V)	UNCLASSIFIED	22278996, 264693
679	76868855 (1357, 1358)	Novel Protein sim. GBank g1382873 (emb1CA02218) - (A1034355) putative ABC transporter (Streptomyces coelicolor)			264600, 264602 264102, 264907, 264908, 265006, 264993, 263972, 33373044, 264586
680	207262424 (1359, 1360)	Novel Protein sim. GBank g15714453 (p1NP_006050) lipi AHC - laminin, gamma 3	Contains protein domain (PF00053) - Laminin EGF-like (Domains III and V)	UNCLASSIFIED	264655 264634
681	11302476 (1363, 1364)	Novel Protein sim. GBank g1758209 (p1NP_004081) lipi AHC - dual specificity phosphatase 3 (Vaccinia virus phosphatase VHI-related)	Contains protein domain (PF00782) - Dual specificity phosphatase, catalytic domain		264605
682	60083680 (1365, 1366)	Novel Protein sim. GBank g15420387 (emb1CA046678) 11 - (A1243459) proteophosphoglycan (Leishmania major)			264909, 263967, 263981 264631 264653, 264635 264907, 264638
683	80246735 (1369, 1370)	Novel Protein sim. GBank g13046330 (AF048330) - PPAR gamma coactivator (Mus musculus)	Contains protein domain (PF00079) - RNA recognition motif (s.a.s. RRM, RBD, or RNP domain)	UNCLASSIFIED	264906, 264907, 265007, 265009, 60433438, 21900754, 264760, 16106358, 21900766, 21900769, 265021, 16106361, 263974, 16106375, 264557, 16106385, 22278902 264510, 264511, 264764, 264769
684	79553412 (1375, 1376)	Novel Protein sim. GBank g12588852 (AF027768) - LspA (Serratia marcescens)			264757
685	80044255 (1377, 1378)	Novel Protein sim. GBank g13046330 (AF048330) - PPAR gamma coactivator (Mus musculus)			55812034, 55811957, 265018, 55811150, 18108351, 264908, 60431628, 264594
686	60386750 (1379, 1380)	Novel Protein sim. GBank g1449684 (p1NP010426) SP62_HUMAN - SPIRICO-SOME ASSOCIATED PROTEIN 62 (SP-62) (SF3068)			264635
687	81854382 (1381, 1382)	Novel Protein sim. GBank g15420387 (emb1CA046678) 11 - (A1243459) proteophosphoglycan (Leishmania major)			
688	83069836 (1383, 1384)	Novel Protein sim. GBank g1354053 (emb1CA046678) 11 - (X53413) U88 [Human herpesvirus 6]			
689	76586116 (1385, 1386)	Novel Protein sim. GBank g1354053 (emb1CA046678) 11 - (X53413) U88 [Human herpesvirus 6]			
690	82455983 (1387, 1388)	Novel Protein sim. GBank g1967327 (p1NP01003) G48_HSVSA - HYPOTHETICAL GENE 48 PROTEIN			22278996, 264510, 264602, 264603, 264762, 264687, 264769, 264693

695	94147649 (1389, 1390)	Novel Protein sim. GBank gl1448339jmc(CAB30059 1) - (A-010901) MUC4 (Homo sapiens)	Contains protein domain (PF00094) - von Willebrand factor type D domain	UNCLASSIFIED	56192575, 264509, 264905, 264907, 29331810, 264909, 264909, 264511, 265007, 264510, 264756, 264784, 264288, 55274791, 264905, 264935
696	79930682 (1391, 1392)	Novel Protein sim. GBank gl26499550 (R301058) - glutamine ABC transporter, ATP-binding protein (GlcA) from <i>Proteobacteria</i> (Bacteria)	Contains protein domain (PF00005) - ABC transporter	transport	264682
697	11767889 (1393, 1394)	Novel Protein sim. GBank gl1731343jmc(Q10894)Y25_MVCTU - HYPOTHETICAL 24.4 KD PROTEIN CY49_25	Contains protein domain (PF01936) - Transposase	UNCLASSIFIED	264682
698	66695862 (1395, 1396)	Novel Protein sim. GBank gl1248933jmc(Q46338)SIXQ_CORS - SARCOSE OXIDASE GAMMA SUBUNIT	Contains protein domain (PF01936) - Transposase	UNCLASSIFIED	264682, 264688, 264693, 264693
699	75452556 (1397, 1398)	Novel Protein sim. GBank gl1001256jmc(BAA10477) - contains similarity to the leishmanin family (Caenorhabditis elegans)	Contains protein domain (PF01936) - Transposase	UNCLASSIFIED	264682, 264688, 264693, 264693
700	75635098 (1399, 1400)	Novel Protein sim. GBank gl1001256jmc(BAA10477) - contains similarity to the leishmanin family (Caenorhabditis elegans)	Contains protein domain (PF01936) - Transposase	UNCLASSIFIED	264682, 264688, 264693, 264693
701	80230242 (1401, 1402)	Novel Protein sim. GBank gl1001256jmc(BAA10477) - contains similarity to the leishmanin family (Caenorhabditis elegans)	Contains protein domain (PF01936) - Transposase	UNCLASSIFIED	264682, 264688, 264693, 264693
702	75817789 (1403, 1404)	Novel Protein sim. GBank gl1248933jmc(Q46338)SIXQ_CORS - SARCOSE OXIDASE GAMMA SUBUNIT	Contains protein domain (PF01936) - Transposase	UNCLASSIFIED	264682, 264688, 264693, 264693
703	20448820 (1405, 1406)	Novel Protein sim. GBank gl1248933jmc(Q46338)SIXQ_CORS - SARCOSE OXIDASE GAMMA SUBUNIT	Contains protein domain (PF01936) - Transposase	UNCLASSIFIED	264682, 264688, 264693, 264693
704	94312224 (1407, 1408)	Novel Protein sim. GBank gl13155513 (AF067219) - contains similarity to the leishmanin family (Caenorhabditis elegans)	Contains protein domain (PF01936) - Transposase	UNCLASSIFIED	264682, 264688, 264693, 264693
705	17932141 (1409, 1410)	Novel Protein sim. GBank gl1421091jmc(J53079) - contains similarity to the leishmanin family (Caenorhabditis elegans)	Contains protein domain (PF01936) - Transposase	UNCLASSIFIED	264682, 264688, 264693, 264693
706	20288062 (1411, 1412)	Novel Protein sim. GBank gl13924872jmc(Q55789)YD7A_SYNY3 - HYPOTHETICAL 52.8 KD PROTEIN SLR007A	Contains protein domain (PF01936) - Transposase	UNCLASSIFIED	264682, 264688, 264693, 264693
707	20633065 (1413, 1414)	Novel Protein sim. GBank gl13420668jmc(AAC31907_1) - (AF075709) ABC transporter ATP-binding subunit (Pseudomonas putida)	Contains protein domain (PF01936) - Transposase	transport	264603
708	20708292 (1415, 1416)	Novel Protein sim. GBank gl13649741jmc(CAA03385) - (A-0000281) mucin (Homo sapiens)	Contains protein domain (PF01936) - Transposase	struct	264607, 264692
709	86001439 (1417, 1418)	Novel Protein sim. GBank gl13080425jmc(CAA18744_1) - (AL027804) putative protein (A. baumannii)	Contains protein domain (PF01936) - Transposase	struct	18103396, 264637, 264908, 264909
710	111356683 (1419, 1420)	Novel Protein sim. GBank gl13080425jmc(CAA18744_1) - (AL027804) putative protein (A. baumannii)	Contains protein domain (PF01936) - Transposase	struct	264369
711	17932141 (1421, 1422)	Novel Protein sim. GBank gl1755686jmc(Q02323) - low density lipoprotein-related protein 1 (alpha-2-macroglobulin receptor)	Contains protein domain (PF00058) - low-density lipoprotein receptor repeat class B	apolipoprotein	265019
712	80255164 (1423, 1424)	Novel Protein sim. GBank gl1755686jmc(Q02323) - low density lipoprotein-related protein 1 (alpha-2-macroglobulin receptor)	Contains protein domain (PF00058) - low-density lipoprotein receptor repeat class B	apolipoprotein	264591
713	78263126 (1425, 1426)	Novel Protein sim. GBank gl1755686jmc(Q02323) - low density lipoprotein-related protein 1 (alpha-2-macroglobulin receptor)	Contains protein domain (PF00058) - low-density lipoprotein receptor repeat class B	hydrolase	264906, 264907
714	27847651 (1427, 1428)	Novel Protein sim. GBank gl1755686jmc(Q02323) - low density lipoprotein-related protein 1 (alpha-2-macroglobulin receptor)	Contains protein domain (PF00058) - low-density lipoprotein receptor repeat class B	hydrolase	264508, 264555

715	79639423 (1429, 1430)	Novel Protein sim. GBank g11790035 (AE000332) - ori, hypothetical protein [<i>Escherichia coli</i>]	UNCLASSIFIED	264907
716	76559072 (1431, 1432)	Novel Protein sim. GBank g12494074ep156533GABD_RHEN - PROBABLE SUCCINATE-SEMIALDEHYDE DEHYDROGENASE (NADP+) (SSDH)	dehydrogenase	264692 264636
717	79491842 (1433, 1434)	Novel Protein sim. GBank g12494074ep156533GABD_RHEN - PROBABLE SUCCINATE-SEMIALDEHYDE DEHYDROGENASE (NADP+) (SSDH)	kinase	18106392, 22276694, 22276936, 265008 265016, 264697, 18106354, 264684, 264685, 18106355, 264698, 264699, 264699, 264699, 18106361, 264691, 264691, 264691, 264691, 264635, 18106381, 18106382, 63373044, 18106386
718	94313658 (1435, 1436)	Novel Protein sim. GBank g13873579emb1CA494661 - (L11173) similar to pro-collagen domains, cDNA EST from <i>Canis familiaris</i> gene, cDNA EST EMBL D27977 comes from this gene, cDNA EST EMBL D34169 comes from this gene, cDNA EST EMBL D64382 comes from this gene, cDNA EST Novel Protein sim. GBank g12194302emb1CA606631 - (295387) hypothetical protein RV2811c [Mycobacterium tuberculosis]	von Willebrand factor type C domain	265011
719	17879564 (1437, 1438)	Novel Protein sim. GBank g12194302emb1CA606631 - (295387) hypothetical protein RV2811c [Mycobacterium tuberculosis]	UNCLASSIFIED	264908
720	79841684 (1439, 1440)	Novel Protein sim. GBank g11235300ep1043291RRPX_PLALO - HISTIDINE-RICH GLYCOPROTEIN PRECURSOR	UNCLASSIFIED	264629
721	15020180 (1441, 1442)	Novel Protein sim. GBank g1498253 (U02372) - integrase [Vibrio cholerae]	UNCLASSIFIED	264910
722	9626503 (1443, 1444)	Novel Protein sim. GBank g1498253 (U02372) - integrase [Vibrio cholerae]	UNCLASSIFIED	264991
723	19156569 (1445, 1446)	Novel Protein sim. GBank g12253054emb1CA510765 - (217295) hypothetical protein RV2114 [Mycobacterium tuberculosis]	protease	264909
724	10125494 (1447, 1448)	Novel Protein sim. GBank g14083015 (AF083061) - protease Pta [Pseudomonas fluorescens]	protease	264909
725	76978679 (1449, 1450)	Novel Protein sim. GBank g12253054emb1CA510765 - (217295) hypothetical protein RV2114 [Mycobacterium tuberculosis]	protease	264909
726	130962622 (1451, 1452)	Novel Protein sim. GBank g12253054emb1CA510765 - (217295) hypothetical protein RV2114 [Mycobacterium tuberculosis]	protease	264909
727	13522872 (1453, 1454)	Novel Protein sim. GBank g12253054emb1CA510765 - (217295) hypothetical protein RV2114 [Mycobacterium tuberculosis]	protease	264909
728	20268471 (1455, 1456)	Novel Protein sim. GBank g12253054emb1CA510765 - (217295) hypothetical protein RV2114 [Mycobacterium tuberculosis]	protease	264909
729	11793753 (1457, 1458)	Novel Protein sim. GBank g12253054emb1CA510765 - (217295) hypothetical protein RV2114 [Mycobacterium tuberculosis]	protease	264909
730	19900373 (1459, 1460)	Novel Protein sim. GBank g12253054emb1CA510765 - (217295) hypothetical protein RV2114 [Mycobacterium tuberculosis]	protease	264909
731	60569750 (1461, 1462)	Novel Protein sim. GBank g12253054emb1CA510765 - (217295) hypothetical protein RV2114 [Mycobacterium tuberculosis]	protease	264909
732	60298175 (1463, 1464)	Novel Protein sim. GBank g12253054emb1CA510765 - (217295) hypothetical protein RV2114 [Mycobacterium tuberculosis]	protease	264909
733	20446339 (1465, 1466)	Novel Protein sim. GBank g12253054emb1CA510765 - (217295) hypothetical protein RV2114 [Mycobacterium tuberculosis]	protease	264909
734	20435687 (1467, 1468)	Novel Protein sim. GBank g12253054emb1CA510765 - (217295) hypothetical protein RV2114 [Mycobacterium tuberculosis]	protease	264909

735	11607559 (1469, 1470)	Novel Protein sim. GBank g1401582sp1p274321YCE_ECOLI - HYPOTHETICAL 48.9 KD PROTEIN IN GLTS-SELIC INTERGENIC REGION	Contains protein domain (PF00528) - Binding-protein-dependent transport systems inner membrane component	264594
738	10879734 (1471, 1472)	Novel Protein sim. GBank g1400831sp1p3113j10POTL_ECOLI - PUTRESCINE TRANSPORT SYSTEM PERMEASE PROTEIN POT1	Contains protein domain (PF00615) - Regulator of G protein signaling domain	264636 265020
737	76945340 (1473, 1474)			265008 264910
738	17895353 (1475, 1476)	Novel Protein sim. GBank g12506867sp1p33225TORA_ECOLI - TRIMETHYLAMINE- N-OXIDE REDUCTASE PRECURSOR (TMAO REDUCTASE) (TRIMETHYLAMINE OXIDASE)	oxidase	264907, 264764, 264634, 264637
739	76933670 (1477, 1478)	Novel Protein sim. GBank g12506867sp1p33225TORA_ECOLI - TRIMETHYLAMINE- N-OXIDE REDUCTASE PRECURSOR (TMAO REDUCTASE) (TRIMETHYLAMINE OXIDASE)	UNCLASSIFIED	264985, 35696286, 264510, 264906, 18108352
740	19981557 (1479, 1480)	Novel Protein sim. GBank g12506867sp1p33225TORA_ECOLI - TRIMETHYLAMINE- N-OXIDE REDUCTASE PRECURSOR (TMAO REDUCTASE) (TRIMETHYLAMINE OXIDASE)	UNCLASSIFIED	264907, 264764, 264634, 264637
741	79827273 (1481, 1482)	Novel Protein sim. GBank g12506867sp1p33225TORA_ECOLI - TRIMETHYLAMINE- N-OXIDE REDUCTASE PRECURSOR (TMAO REDUCTASE) (TRIMETHYLAMINE OXIDASE)	Domain of unknown function	264985, 35696286, 264510, 264906, 18108352
742	82393792 (1483, 1484)	Novel Protein sim. GBank g12506867sp1p33225TORA_ECOLI - TRIMETHYLAMINE- N-OXIDE REDUCTASE PRECURSOR (TMAO REDUCTASE) (TRIMETHYLAMINE OXIDASE)	UNCLASSIFIED	264907, 264764, 264634, 264637
743	82300051 (1485, 1486)	Novel Protein sim. GBank g12506867sp1p33225TORA_ECOLI - TRIMETHYLAMINE- N-OXIDE REDUCTASE PRECURSOR (TMAO REDUCTASE) (TRIMETHYLAMINE OXIDASE)	UNCLASSIFIED	264907, 264764, 264634, 264637
744	80230421 (1487, 1488)	Novel Protein sim. GBank g12506867sp1p33225TORA_ECOLI - TRIMETHYLAMINE- N-OXIDE REDUCTASE PRECURSOR (TMAO REDUCTASE) (TRIMETHYLAMINE OXIDASE)	UNCLASSIFIED	264907, 264764, 264634, 264637
745	9841963 (1489, 1490)	Novel Protein sim. GBank g12506867sp1p33225TORA_ECOLI - TRIMETHYLAMINE- N-OXIDE REDUCTASE PRECURSOR (TMAO REDUCTASE) (TRIMETHYLAMINE OXIDASE)	UNCLASSIFIED	264907, 264764, 264634, 264637
746	11073229 (1491, 1492)	Novel Protein sim. GBank g12506867sp1p33225TORA_ECOLI - TRIMETHYLAMINE- N-OXIDE REDUCTASE PRECURSOR (TMAO REDUCTASE) (TRIMETHYLAMINE OXIDASE)	UNCLASSIFIED	264907, 264764, 264634, 264637
747	94322044 (1493, 1494)	Novel Protein sim. GBank g12506867sp1p33225TORA_ECOLI - TRIMETHYLAMINE- N-OXIDE REDUCTASE PRECURSOR (TMAO REDUCTASE) (TRIMETHYLAMINE OXIDASE)	UNCLASSIFIED	264907, 264764, 264634, 264637
748	11617823 (1495, 1496)	Novel Protein sim. GBank g12506867sp1p33225TORA_ECOLI - TRIMETHYLAMINE- N-OXIDE REDUCTASE PRECURSOR (TMAO REDUCTASE) (TRIMETHYLAMINE OXIDASE)	UNCLASSIFIED	264907, 264764, 264634, 264637

749	20469119 (1497, 1498)	Novel Protein sim. GBank g11169722pP44946PFC_HAELN. FORMAMIDOPYRIMIDINE-DNA GLYCOSYLASE (FAPY-DNA GLYCOSYLASE)	UNCLASSIFIED	264604 264600
750	20296427 (1498, 1500)	Novel Protein sim. GBank g1536068g1A04D251.1AT*15968 - (AF15968) serine/threonine kinase PKN3 [Mycrococcus xanthus]	kinase	264605, 264559
751	21636169 (1501, 1502)	Novel Protein sim. GBank g1116682pP44428[BIOA_HAELN - ADENOSYLMETHIONINE-6-AMINO-7-OXONONANOATE AMINO TRANSFERASE (7,8-DIAMINO-PELAGONIC ACID AMINO TRANSFERASE) (DAPA AMINOTRANSFERASE)	UNCLASSIFIED	264538, 264507, 264510, 265011, 264762, 264689, 35486852, 264638, 16106387
752	82450366 (1503, 1504)	Novel Protein sim. GBank g12851509pP32399YHGE_BACSU - HYPOTHETICAL PROTEIN IN HEM-GLT INTERGENIC REGION (GRB)	UNCLASSIFIED	264909, 264600, 264602, 264604, 264760, 264760, 264634
753	60503718 (1505, 1506)	Novel Protein sim. GBank g12851509pP32399YHGE_BACSU - HYPOTHETICAL PROTEIN IN HEM-GLT INTERGENIC REGION (GRB)	UNCLASSIFIED	264506, 264606, 264607, 264608, 264909, 264759, 264602, 264764, 264760, 264628, 264629, 264630, 264632, 264634, 264635, 264637, 264638, 83373044, 16106386
754	55033741 (1507, 1508)	Novel Protein sim. GBank g13449276jemb(CAA20429) - (A1031317) putative dehydrogenase [Streptomyces coelicolor]	UNCLASSIFIED	264446, 264590 264769, 264586, 264638, 264639
755	50185448 (1500, 1510)	Novel Protein sim. GBank g13449276jemb(CAA20429) - (A1031317) putative dehydrogenase [Streptomyces coelicolor]	UNCLASSIFIED	264446, 264590 264769, 264586, 264638, 264639
756	94631666 (1511, 1512)	Novel Protein sim. GBank g13449276jemb(CAA20429) - (A1031317) putative dehydrogenase [Streptomyces coelicolor]	UNCLASSIFIED	264446, 264590 264769, 264586, 264638, 264639
757	79468533 (1513, 1514)	Novel Protein sim. GBank g13449276jemb(CAA20429) - (A1031317) putative dehydrogenase [Streptomyces coelicolor]	UNCLASSIFIED	264446, 264590 264769, 264586, 264638, 264639
758	76953176 (1515, 1516)	Novel Protein sim. GBank g13449276jemb(CAA20429) - (A1031317) putative dehydrogenase [Streptomyces coelicolor]	UNCLASSIFIED	264446, 264590 264769, 264586, 264638, 264639
759	78475667 (1517, 1518)	Novel Protein sim. GBank g13449276jemb(CAA20429) - (A1031317) putative dehydrogenase [Streptomyces coelicolor]	UNCLASSIFIED	264446, 264590 264769, 264586, 264638, 264639
760	87628888 (1519, 1520)	Novel Protein sim. GBank g13449276jemb(CAA20429) - (A1031317) putative dehydrogenase [Streptomyces coelicolor]	UNCLASSIFIED	264446, 264590 264769, 264586, 264638, 264639
761	79877966 (1521, 1522)	Novel Protein sim. GBank g13449276jemb(CAA20429) - (A1031317) putative dehydrogenase [Streptomyces coelicolor]	UNCLASSIFIED	264446, 264590 264769, 264586, 264638, 264639
762	60023563 (1523, 1524)	Novel Protein sim. GBank g13449276jemb(CAA20429) - (A1031317) putative dehydrogenase [Streptomyces coelicolor]	UNCLASSIFIED	264446, 264590 264769, 264586, 264638, 264639
763	20294813 (1525, 1526)	Novel Protein sim. GBank g13449276jemb(CAA20429) - (A1031317) putative dehydrogenase [Streptomyces coelicolor]	UNCLASSIFIED	264446, 264590 264769, 264586, 264638, 264639
764	38515024 (1527, 1528)	Novel Protein sim. GBank g13449276jemb(CAA20429) - (A1031317) putative dehydrogenase [Streptomyces coelicolor]	UNCLASSIFIED	264446, 264590 264769, 264586, 264638, 264639

768	80507844 (1575, 1576)	Novel Protein sim. GBank gl2746579 (AF015310) - BTH1 (Brassic napus)	synthase	254909, 264602, 264603, 264769, 264638
769	17294715 (1577, 1578)	Novel Protein sim. GBank gl2351149 (U93557) - 40 Cba heat shock chaperone protein (Hsp90alpha cutanobium)	UNCLASSIFIED	265007
790	86284406 (1579, 1580)	Novel Protein sim. GBank gl5706576 (B0A83099.1) - (AB026118) MAL T1 (Homo sapiens)	glycoprotein	2278999, 26331824, 264828, 87168559, 265018, 21906765, 21906767, 21906768, 21906769, 265020, 264692, 22779000, 264963
791	94651627 (1581, 1582)	Novel Protein sim. GBank gl5689924 (emb CA851985.1 - (AL109663) putative isochoRNA synthetase	struct	264601, 264605, 264638
792	80056786 (1583, 1584)	Novel Protein sim. GBank gl393194 (U02379) - S antigen (Phasmodium falciparum)	struct	265021, 264631, 264635, 264556
793	75638730 (1585, 1586)	Novel Protein sim. GBank gl1345408 (b b BAA05046 - (U00001) DnaK (Homo sapiens)	homodimer	264693
794	81839294 (1587, 1588)	Novel Protein sim. GBank gl10585494 (gl S24023 deapamine receptor D4 - human fragment)	UNCLASSIFIED	264693, 264699, 264757
795	80074888 (1589, 1590)	Novel Protein sim. GBank gl1877334 (emb CA807082 - (Z92771) bta (Mycoplasma tuberculosis)	carboxylase	2654448, 35656632, 264502, 264807, 265010, 35656423, 264636
796	86569451 (1591, 1592)	Novel Protein sim. GBank gl2995447 (emb CA971519 - (Y10495) CDV-1R protein (Mus musculus)	struct	60432279, 56811150, 264630, 264637, 264565
797	87771781 (1593, 1594)	Novel Protein sim. GBank gl2995447 (emb CA971519 - (Y10495) CDV-1R protein (Mus musculus)	struct	2278998, 264093, 264094, 56714117, 21906767, 21906769, 265020, 265022
799	79855209 (1595, 1596)	Novel Protein sim. GBank gl4647250 (emb CA831575 - (AL035568) probable Gu-RNA Gln amidotransferase subunit (Streptococcus colocol)	transcriptase	264687, 264768, 264693
800	75927839 (1599, 1600)	Novel Protein sim. GBank gl2791517 (emb CA616564 - (U0321249) hypothetical protein R62477 (Mycobacterium tuberculosis)	hydrolase	264909, 264910, 264636, 264638
801	80499399 (1601, 1602)	Novel Protein sim. GBank gl4887211 (pp A0323237_1 AF14744_ (AF147449) penicillin binding protein 1B (Pseudomonas aeruginosa)	UNCLASSIFIED	264468
802	75834498 (1603, 1604)	Novel Protein sim. GBank gl1176152 (sp P44507 VHA0_HAEN - HYPOTHETICAL PROTEIN H80091	transport	264909, 264511, 265006, 265009, 264769, 264907, 264486
803	20467520 (1605, 1606)	Novel Protein sim. GBank	struct	264605
804	10174239 (1607, 1608)	Novel Protein sim. GBank	kinase	264510
805	75599993 (1609, 1610)	Novel Protein sim. GBank gl2764671 (emb CA640683 - (AJ001330) ornithine transcarbamoylase (Lactobacillus sakei)	transferase	264508
806	80484113 (1611, 1612)	Novel Protein sim. GBank	transferase	264769
807	86381612 (1613, 1614)	Novel Protein sim. GBank gl6833311 (pp 0216201 TFD_CAEEL - HYPOTHETICAL	carbamoyltransferase	264764

808	35106817 (1615, 1616)	Novel Protein sim. GBank gi5913092[gi5913092]CLOPE - ARGININE/CORNITHINE ANTI-PORTER		transport	264909, 264602, 21908764, 18 08374
809	81454254 (1617, 1618)	Novel Protein sim. GBank gi3913019[gi3913019]ALF1, SYNY3 - FRUCTOSE- BISPHOSPHATE ALDOLASE CLASS 1 [FBP ALDOLASE]	Contains protein domain (PF00274) - Fructose-bisphosphate aldolase class-1	UNCLASSIFIED	264508, 264006, 264009, 265007, 264910, 264758, 264800, 264602, 264603, 264806, 264607, 264789, 264693, 264636, 264486
810	86192761 (1619, 1620)	Novel Protein sim. GBank gi401472[gi401472]VAFB, ECOLI - HYPOTHETICAL OXIDOREDUCTASE IN ASPERMATID INTERGENIC REGION	AdoKeto reductase family	reductase	264359
811	80076280 (1621, 1622)			UNCLASSIFIED	264558
812	10297654 (1623, 1624)			UNCLASSIFIED	264632
813	79812280 (1625, 1626)			UNCLASSIFIED	264906
814	80473427 (1627, 1628)			synthase	264789, 264636
815	95419513 (1629, 1630)	Novel Protein sim. GBank gi146168 [U01617] - glutamyl- trycA synthetase [Escherichia coli] Novel Protein sim. GBank gi459652[gi459652] - [P082322] ADAM1004 protein [Homo sapiens]		UNCLASSIFIED	264488, 22278698, 22278699, 29331822, 29331824, 29331825, 29331827, 29331828, 29146495, 264905, 264908, 265007, 3367102, 69433395, 69433435, 264758, 264759, 264760, 264761, 264762, 264763, 264764, 264805, 21908765, 21908767, 265020, 265021, 264692, 65274620, 33671103, 264629, 18108378, 284635, 264638, 60170394, 56182323, 264664
816	19881950 (1631, 1632)	Novel Protein sim. GBank gi178114[embi]CA06254] - (Z03865) hypothetical protein RV0089 [Mycobacterium tuberculosis]		UNCLASSIFIED	264595
817	95293316 (1633, 1634)				
818	90351930 (1635, 1636)	Novel Protein sim. GBank gi14177468 [U3244] - vacuolar protein sorting homolog v-pa33a [Rattus norvegicus]			85274572, 22278698, 60424269, 35665052, 55612038, 21908768, 55611957, 35665917, 33657023, 21908370, 18108374, 18108375, 3569423, 55811576, 284636
819	80254977 (1637, 1638)	Novel Protein sim. GBank gi1001352[gi1001352] - (D04059) ABC class 1 protein [Drosophila sp.]		transport	264555
820	80059688 (1639, 1640)	Novel Protein sim. GBank gi59681[gi59681]YBRT, BACSU - HYPOTHETICAL 74.3 KD PROTEIN IN RPL-COTF INTERGENIC REGION		UNCLASSIFIED	264600, 364602, 264604
821	79762590 (1641, 1642)			UNCLASSIFIED	
822	80215310 (1643, 1644)	Novel Protein sim. GBank gi387800[embi]CA05828] - (Z1264) predicted using GeneIndex: Weak similarity to Mouse T-complex-associated-testes-expressed-1 protein (PIR Acc. No. A45841): cDNA EST EMBL:D32742 comes from this gene; cDNA EST EMBL:D33817 comes from this gene; cDNA EST...		UNCLASSIFIED	264910
823	94922269 (1645, 1646)			stud	264510, 264504, 264637, 264509, 264687, 264681
824	80411171 (1647, 1648)	Novel Protein sim. GBank gi1370079[embi]CA066887] - (X98235) type I [Drosophila melanogaster]	Contains protein domain (PF01429) - Methyl-CpG binding domain		264910, 264163, 264785, 264789, 264683

825	20539600 (1649, 1650)	Novel Protein sim. GBank gi3025132 ep77391 TEAC_ECOLI - HYPOTHETICAL 74.5 KD PROTEIN IN CAPARND INTERGENIC REGION	UNCLASSIFIED	264592
826	11073047 (1651, 1652)	Novel Protein sim. GBank gi324243 tm00A10669 - (A021648) Hypothetical protein RV2202 [Mycobacterium tuberculosis]		264605
827	80054207 (1653, 1654)	Novel Protein sim. GBank gi3417424 tm5 CA020312 - (A031261) putative transport protein (Schistosoma carunculosa pombel)		264603
828	95106322 (1655, 1656)	Novel Protein sim. GBank gi4336692 gi AA017897 - (AF101361) Abnormal X segregation [Drosophila melanogaster]	UNCLASSIFIED	264590, 264508, 264905, 264509, 264906, 264907, 264908, 264909, 264512, 264910, 264591, 264756, 264600, 264766, 264768, 21906768, 35695917, 264691, 264628, 264630, 264631, 264632, 264634, 264635, 264636, 264637, 264658, 264639, 83373044, 18100385, 264663, 264596, 264486
829	81742215 (1657, 1658)	Novel Protein sim. GBank gi3346039 (AF108002) - UOP-N- acetylneuramyl lipase synthetase MurC [Helicobacter mobilis]	UNCLASSIFIED	264758, 264534
830	20396091 (1659, 1660)	Novel Protein sim. GBank gi1870594 tm0CA060553 - (292359) Hypothetical protein RV1024 [Mycobacterium tuberculosis]	UNCLASSIFIED	264603
831	87112435 (1661, 1662)	Novel Protein sim. GBank	UNCLASSIFIED	66714117, 264910, 264639
832	15535322 (1663, 1664)	Novel Protein sim. GBank gi1870594 tm0CA060553 - (292359) Hypothetical protein RV1024 [Mycobacterium tuberculosis]	UNCLASSIFIED	264906
833	20726654 (1665, 1666)	Novel Protein sim. GBank gi2500058 ep Q46287 PFLA_CLOPA - PYRUVATE FORMATE-LYASE ACTIVATING ENZYME	UNCLASSIFIED	264602
834	21429762 (1667, 1668)	Novel Protein sim. GBank gi4947531 ep Q46078 KPYK_CORGL - PYRUVATE KINASE (PK)	kinase	264600, 264602, 264766, 264689, 264636
835	94140482 (1669, 1670)			
836	66126552 (1671, 1672)	Novel Protein sim. GBank gi699315 (U15164) - phosphate transport protein PSTC [Mycobacterium leprae]	transport	264765, 263994, 21906767, 264910, 264632, 264635, 264259, 264639, 264653, 83373044, 264758, 39690052, 22279002, 264500, 264905, 264906, 264448, 263972, 264909, 264909
837	70452450 (1673, 1674)	Novel Protein sim. GBank	UNCLASSIFIED	35695917, 264557
838	70184203 (1675, 1676)	gi 2886776 ep Q0502 APG_ARATH - INTER-SPECIFIC PROLINE-RICH PROTEIN APC PRECURSOR	UNCLASSIFIED	264595
839	76841125 (1677, 1678)	Novel Protein sim. GBank gi2496533 ep Q50598 YOD8_MYCTU - HYPOTHETICAL 69.9 KD PROTEIN CY1A11.00	UNCLASSIFIED	264607
840	80059851 (1679, 1680)	Novel Protein sim. GBank gi4557753 ep NP_000372 1pMID1 - midline 1 protein (liver)	UNCLASSIFIED	264906
		Contains protein domain (PF00097) - Zinc finger, C3HC4 type (RING finger)	interferon receptor	264762, 264556

853	13032587 (1705, 1706)	Novel Protein sim. GBank gi340236(emb)(CA476082) - (Y16136) 2-oxoalate reductase (Morella thermacetica)	reductase	2646336 2645668
854	80532438 (1707, 1708)	Novel Protein sim. GBank gi340236(emb)(CA476082) - (Y16136) 2-oxoalate reductase (Morella thermacetica)	reductase	2646336 2645668
855	78841130 (1709, 1710)	Novel Protein sim. GBank gi340236(emb)(CA476082) - (Y16136) 2-oxoalate reductase (Morella thermacetica)	reductase	2646336 2645668
856	13942436 (1711, 1712)	Novel Protein sim. GBank gi340236(emb)(CA476082) - (Y16136) 2-oxoalate reductase (Morella thermacetica)	reductase	2646336 2645668
857	12410103 (1713, 1714)	Novel Protein sim. GBank gi340236(emb)(CA476082) - (Y16136) 2-oxoalate reductase (Morella thermacetica)	reductase	2646336 2645668
858	80246910 (1715, 1716)	Novel Protein sim. GBank gi340236(emb)(CA476082) - (Y16136) 2-oxoalate reductase (Morella thermacetica)	reductase	2646336 2645668
859	20246634 (1717, 1718)	Novel Protein sim. GBank gi340236(emb)(CA476082) - (Y16136) 2-oxoalate reductase (Morella thermacetica)	reductase	2646336 2645668
860	80041749 (1719, 1720)	Novel Protein sim. GBank gi340236(emb)(CA476082) - (Y16136) 2-oxoalate reductase (Morella thermacetica)	reductase	2646336 2645668
861	65557045 (1721, 1722)	Novel Protein sim. GBank gi340236(emb)(CA476082) - (Y16136) 2-oxoalate reductase (Morella thermacetica)	reductase	2646336 2645668
862	80079467 (1723, 1724)	Novel Protein sim. GBank gi340236(emb)(CA476082) - (Y16136) 2-oxoalate reductase (Morella thermacetica)	reductase	2646336 2645668
863	80579931 (1725, 1726)	Novel Protein sim. GBank gi340236(emb)(CA476082) - (Y16136) 2-oxoalate reductase (Morella thermacetica)	reductase	2646336 2645668
864	94339904 (1727, 1728)	Novel Protein sim. GBank gi340236(emb)(CA476082) - (Y16136) 2-oxoalate reductase (Morella thermacetica)	reductase	2646336 2645668
865	80046310 (1729, 1730)	Novel Protein sim. GBank gi340236(emb)(CA476082) - (Y16136) 2-oxoalate reductase (Morella thermacetica)	reductase	2646336 2645668
866	80162031 (1731, 1732)	Novel Protein sim. GBank gi340236(emb)(CA476082) - (Y16136) 2-oxoalate reductase (Morella thermacetica)	reductase	2646336 2645668
867	80052402 (1733, 1734)	Novel Protein sim. GBank gi340236(emb)(CA476082) - (Y16136) 2-oxoalate reductase (Morella thermacetica)	reductase	2646336 2645668
868	10075384 (1735, 1736)	Novel Protein sim. GBank gi340236(emb)(CA476082) - (Y16136) 2-oxoalate reductase (Morella thermacetica)	reductase	2646336 2645668
869	80052406 (1737, 1738)	Novel Protein sim. GBank gi340236(emb)(CA476082) - (Y16136) 2-oxoalate reductase (Morella thermacetica)	reductase	2646336 2645668
870	80246951 (1739, 1740)	Novel Protein sim. GBank gi340236(emb)(CA476082) - (Y16136) 2-oxoalate reductase (Morella thermacetica)	reductase	2646336 2645668
871	20378295 (1741, 1742)	Novel Protein sim. GBank gi340236(emb)(CA476082) - (Y16136) 2-oxoalate reductase (Morella thermacetica)	reductase	2646336 2645668
872	95197114 (1743, 1744)	Novel Protein sim. GBank gi340236(emb)(CA476082) - (Y16136) 2-oxoalate reductase (Morella thermacetica)	reductase	2646336 2645668
873	20189728 (1745, 1746)	Novel Protein sim. GBank gi340236(emb)(CA476082) - (Y16136) 2-oxoalate reductase (Morella thermacetica)	reductase	2646336 2645668

874	80071692 (1747, 1748)	Novel Protein sim. GBank (1747) HYPHORE, BACSUS, SUCROSE & PHOSPHATASE (HYPHORE, BACSUS, SUCROSE & PHOSPHATASE)		UNCLASSIFIED	264600
875	86080446 (1748, 1750)	Novel Protein sim. GBank g1481000 (p153754 - mucin - human (fragment))			264259, 264448, 264288, 264557, 87168518
876	86465157 (1751, 1752)	Novel Protein sim. GBank g1328283 (A5010460) - iron(III) dicluate transport ATP-binding protein [Rhodococcus capsulatus]	Contains protein domain (PF00005) - transport ABC transporter		264907, 264601, 264602, 264605, 265020, 60431602
877	87802248 (1753, 1754)	Novel Protein sim. GBank g1731074 (p140349) (URB1 - SIDEROPHORE BIOSYNTHESIS REGULATORY PROTEIN URB1)	Contains protein domain (PF00320) - transcript factor GATA zinc finger		22278998, 264906, 264369
878	80187286 (1755, 1756)	Novel Protein sim. GBank g11351614 (p1009853) (YAC - SIDEROPHORE BIOSYNTHESIS REGULATORY PROTEIN URB1)	ATPase associated		264389, 264555
879	94328662 (1757, 1758)	Novel Protein sim. GBank g13875304 (p1009853) (YAC - SIDEROPHORE BIOSYNTHESIS REGULATORY PROTEIN URB1)			56182575, 26531824, 264508, 264906, 265018, 18106351, 264448, 264683, 21960788, 21960788, 60710615, 33857023, 56274620, 33857103, 18106374, 35695655, 264503
880	8491135 (1759, 1760)	Novel Protein sim. GBank g1137120 (p11214) (HUR0T, MOUSE - TISSUE PLASMINOGEN ACTIVATOR PRECURSOR (TPA) (T- PLASMINOGEN ACTIVATOR))	Contains protein domain (PF00051) - cathepsin Kringle domain		264508
881	11290122 (1761, 1762)	Novel Protein sim. GBank g1263208 (p1009853) (YAC - SIDEROPHORE BIOSYNTHESIS REGULATORY PROTEIN URB1)		UNCLASSIFIED	264508
882	11077011 (1763, 1764)	Novel Protein sim. GBank g1263208 (p1009853) (YAC - SIDEROPHORE BIOSYNTHESIS REGULATORY PROTEIN URB1)		UNCLASSIFIED	264556
883	79582968 (1765, 1766)	Novel Protein sim. GBank g1155068 (p1009853) (YAC - SIDEROPHORE BIOSYNTHESIS REGULATORY PROTEIN URB1)		UNCLASSIFIED	264688
884	13517921 (1767, 1768)	Novel Protein sim. GBank g12078027 (p1009853) (YAC - SIDEROPHORE BIOSYNTHESIS REGULATORY PROTEIN URB1)		UNCLASSIFIED	264636
885	9052457 (1769, 1770)	Novel Protein sim. GBank g12078027 (p1009853) (YAC - SIDEROPHORE BIOSYNTHESIS REGULATORY PROTEIN URB1)		UNCLASSIFIED	264605, 18106362
886	11685706 (1771, 1772)	Novel Protein sim. GBank g1263208 (p1009853) (YAC - SIDEROPHORE BIOSYNTHESIS REGULATORY PROTEIN URB1)			264600
887	94315307 (1773, 1774)	Novel Protein sim. GBank g1263208 (p1009853) (YAC - SIDEROPHORE BIOSYNTHESIS REGULATORY PROTEIN URB1)	dehydrogenase		33569603, 264605, 264600, 264603, 33569597, 33569585, 264638
888	10083398 (1775, 1776)	Novel Protein sim. GBank g1263208 (p1009853) (YAC - SIDEROPHORE BIOSYNTHESIS REGULATORY PROTEIN URB1)	UNCLASSIFIED		264608
889	20385917 (1777, 1778)	Novel Protein sim. GBank g1263208 (p1009853) (YAC - SIDEROPHORE BIOSYNTHESIS REGULATORY PROTEIN URB1)			264603
890	19904337 (1779, 1780)	Novel Protein sim. GBank g1263208 (p1009853) (YAC - SIDEROPHORE BIOSYNTHESIS REGULATORY PROTEIN URB1)			264629

891	13516073 (1781, 1782)	Novel Protein sim. GBank g14659369p10AD34331.1(AF-11248) RAD54B protein (Homo sapiens)	UNCLASSIFIED	264636
892	87634157 (1783, 1784)	Novel Protein sim. GBank g1545526b143333 - LEP- 1b-transcription factor binding to initiation site of HIV-1 (alternatively spliced) [human, Namalwa cells, Peptide, 541 aa]	transcription factor	22278968, 22278969, 26331828, 35690052, 264908, 264909, 265008, 265011, 264602, 265019, 264766, 21906765, 21906766, 21906769, 265020, 265021, 56526486 264689, 265367
893	79168037 (1785, 1786)	Novel Protein sim. GBank g12326689p10B0608CYSK_MAZE - CYSTEINE SYNTHASE (O-ACETYL SERINE SULFHYDRYLASE) (O- ACETYL SERINE (THIO-L-LYASE) (CSASE)	synthase	
894	11102240 (1787, 1788)		UNCLASSIFIED	263978
895	50238585 (1789, 1790)		UNCLASSIFIED	264632
896	79374163 (1791, 1792)		UNCLASSIFIED	264896, 26331828, 264511
897	94991923 (1793, 1794)		UNCLASSIFIED	55811957, 22779000, 264466
898	87855109 (1795, 1796)		UNCLASSIFIED	264631
899	11100463 (1797, 1798)	Novel Protein sim. GBank g11750127 (U64460) - YncC [Bacillus subtilis]	transport	264769, 264691, 264563
900	80496788 (1799, 1800)	Novel Protein sim. GBank g13122879p10074393SYA_MYCTU - ALANYL-TRNA SYNTHETASE (ALANINE-TRNA LIGASE) (ALARS)	synthase	264907, 264602, 264603, 264769, 35695917, 18106376, 264563
901	80502410 (1801, 1802)	Novel Protein sim. GBank g13355707 (emb)CAA200011 - (ALD031124) 3-isopropylmalate dehydratase large subunit (Stereomyces coelicolor)	isomerase	264909, 265008, 264602, 264804, 264769, 264689, 265493
902	80503301 (1803, 1804)	Novel Protein sim. GBank g12560120 (emb)CAA18018.11 - (ALD02121) glbK (Mycobacterium tuberculosis)	kinase	35696052, 264903, 264511, 264512, 264605, 264760, 18100351, 264762, 264687, 264768, 264769, 264688, 21900784, 35695917, 27486262, 35695853, 264634, 264836, 264486
903	82660206 (1805, 1806)	Novel Protein sim. GBank g12560120 (emb)CAA18018.11 - (ALD02121) glbK (Mycobacterium tuberculosis)	kinase	264604
904	20451078 (1807, 1808)	Novel Protein sim. GBank g12738832p10A0900A9GL_COCM - ARGININASE	hydrolase	264909
905	9398443 (1809, 1810)	Novel Protein sim. GBank g14567200p10A022616.1 (AC00716 - (AC007168) hypothetical protein (Arabidopsis thaliana))	UNCLASSIFIED	264909
906	80056283 (1811, 1812)	Novel Protein sim. GBank g13782002 (emb)CAA212621 - (ALD031855) putative vacuolar membrane protein (Schizosaccharomyces pombe)	UNCLASSIFIED	264909
907	87913201 (1813, 1814)	Novel Protein sim. GBank g1366544 (doi)BA342361 - (AB015054) Arg2 [Rhizomucor pusillus]	UNCLASSIFIED	264909
908	87913201 (1813, 1814)	Novel Protein sim. GBank g1366544 (doi)BA342361 - (AB015054) Arg2 [Rhizomucor pusillus]	UNCLASSIFIED	264909
909	20727807 (1817, 1818)	Novel Protein sim. GBank g14568720 (doi)BA476883.11 - (AB003137) DnaJ homolog protein (Salix glauca)	UNCLASSIFIED	264602
910	16776206 (1819, 1820)	Novel Protein sim. GBank g14568720 (doi)BA476883.11 - (AB003137) DnaJ homolog protein (Salix glauca)	eph	265009

963	80590374 (1925, 1926)				UNCLASSIFIED	264510, 264288, 264555, 264556, 264557, 264486
964	79832019 (1927, 1928)	Novel Protein sim. GBank g14-958622[db][BAAF6833.1] - (AB023206) KIAA0099 protein (Homo sapiens)			UNCLASSIFIED	264112, 264910, 264689
965	91239445 (1929, 1930)	Novel Protein sim. GBank g12-62387[emb][CA46678.1] - (U2503429) proteoglycan[Canis familiaris major]			UNCLASSIFIED	264488, 265017, 264443, 264634, 264558, 263758, 264892, 264596, 264604, 264789, 21903794, 264892, 264903, 264629, 264638, 264639
966	93292815 (1931, 1932)				phosphate	264780
967	79235708 (1933, 1934)	Novel Protein sim. GBank g117312076[db][1156]RGX3 MYCTU - SENSORY TRANSDUCTION PROTEIN RGX3			Response regulator receiver domain	
968	79560269 (1935, 1936)	Novel Protein sim. GBank g12-66183[emb][CAA75187.1] - (Y14864) putative transport protein [Methylophilus methylophilus]			transport	264693
969	79919470 (1937, 1938)	Novel Protein sim. GBank g1541887[emb][CA46422.1] - (AL086747) hypothetical protein [Homo sapiens]			Contains protein domain (PF00089) - zinc finger, C2H2 type	35598286, 264685, 264686, 35595917, 264692, 18108374, 264635
970	95059547 (1939, 1940)				UNCLASSIFIED	18108392, 18108394, 18108398, 22278996, 22278998, 22278999, 29147520, 264828, 265006, 265007, 265008, 265009, 16103346, 33109854, 265010, 265011, 16103351, 264285, 21908797, 21908798, 16103370, 16103371, 16103372, 264635, 18108396, 83373044, 18108397, 18108398
971	78919770 (1941, 1942)				UNCLASSIFIED	265007, 265020, 22278902
972	20710704 (1943, 1944)					264557
973	20370183 (1945, 1946)	Novel Protein sim. GBank g11723119[db][P33990]Y174_HUMAN - HYPOTHETICAL PROTEIN KIAA0174				264604
974	80057103 (1947, 1948)				UNCLASSIFIED	264555
975	10186018 (1949, 1950)				UNCLASSIFIED	264510
976	80205742 (1951, 1952)	Novel Protein sim. GBank g13-88145[emb][CA62988.1] - (Z89753) predicted using Genfinder. Similarity to Yeast Hypothetical protein Y19 (SWY19_YEAST): cDNA EST EMBL:027680 comes from the gene: cDNA EST EMBL:027679 comes from the gene: cDNA EST EMBL:027677 comes from the ge...			UNCLASSIFIED	264508, 264906, 264758, 264532, 264639, 264563
977	10355349 (1953, 1954)	Novel Protein sim. GBank g15-44456[db][Q5533MY33_PSEPI - XYLDLGF OPERON TRANSCRIPTIONAL ACTIVATOR 3			UNCLASSIFIED	264906
978	80022592 (1955, 1956)				UNCLASSIFIED	264600, 264602, 264603, 264604
979	80447820 (1957, 1958)	Novel Protein sim. GBank g13-7190[emb][CAA75688] - (Y15909) DIA-12C protein [Homo sapiens]			UNCLASSIFIED	264767, 264768, 265008, 265007, 264906
980	80022592 (1959, 1960)				UNCLASSIFIED	264600, 264602, 264605
981	80098550 (1961, 1962)	Novel Protein sim. GBank g13-59594-0 (A-07389) - (F03081) bicogential dysplasia protein 2 [Mus musculus]			UNCLASSIFIED	264692, 264555, 264556, 264557, 264559

50195570 (1963, 1964)	Novel Protein sm. GBank g12650220(temb)CA0715791 - (Y10545) fused-cds5 [Eschrichia coli]	UNCLASSIFIED	264404
90956047 (1965, 1966)	Novel Protein sm. GBank g1476393(ajl)B34302 - myosin heavy chain-B, neuronal - chicken	struct	65274572, 55182575, 264908, 264909, 265007, 265008, 264758, 265010, 58811150, 33557023, 264034, 264547, 264553
20466878 (1967, 1968)	Novel Protein sm. GBank g1345150(temb)CA0766011 - (A007747) hypothetical protein BLSP1.21 [Bordetella bronchiseptica]	UNCLASSIFIED	264605
55451568 (1969, 1970)	Novel Protein sm. GBank g12455056(ajl)AA076801 - g12725506(ajl)AA084800.1 [AF060152] MET1H protein [Homo sapiens]	transferase	56182435, 264600
87102868 (1971, 1972)	Novel Protein sm. GBank g12650220(temb)CA0715791 - (Y10545) fused-cds5 [Eschrichia coli]	UNCLASSIFIED	264106, 264110, 265020, 60770615
76857231 (1973, 1974)	Novel Protein sm. GBank g12650220(temb)CA0715791 - (Y10545) fused-cds5 [Eschrichia coli]	UNCLASSIFIED	264609
19859601 (1975, 1976)	Novel Protein sm. GBank g12650220(temb)CA0715791 - (Y10545) fused-cds5 [Eschrichia coli]	UNCLASSIFIED	264600
86095329 (1977, 1978)	Novel Protein sm. GBank g12650220(temb)CA0715791 - (Y10545) fused-cds5 [Eschrichia coli]	UNCLASSIFIED	264600
86057746 (1978, 1980)	Novel Protein sm. GBank g12650220(temb)CA0715791 - (Y10545) fused-cds5 [Eschrichia coli]	UNCLASSIFIED	264600
10106810 (1981, 1982)	Novel Protein sm. GBank g12650220(temb)CA0715791 - (Y10545) fused-cds5 [Eschrichia coli]	UNCLASSIFIED	264600
76845694 (1983, 1984)	Novel Protein sm. GBank g12650220(temb)CA0715791 - (Y10545) fused-cds5 [Eschrichia coli]	UNCLASSIFIED	264600
10814053 (1985, 1986)	Novel Protein sm. GBank g12650220(temb)CA0715791 - (Y10545) fused-cds5 [Eschrichia coli]	UNCLASSIFIED	264600
11090590 (1987, 1988)	Novel Protein sm. GBank g12650220(temb)CA0715791 - (Y10545) fused-cds5 [Eschrichia coli]	UNCLASSIFIED	264600
94321911 (1989, 1990)	Novel Protein sm. GBank g12650220(temb)CA0715791 - (Y10545) fused-cds5 [Eschrichia coli]	UNCLASSIFIED	264600
915106572(ajl)AA039760.1 (AF14384 - (AF14394) transcriptional activator SRCAP [Homo sapiens])	Novel Protein sm. GBank g12650220(temb)CA0715791 - (Y10545) fused-cds5 [Eschrichia coli]	UNCLASSIFIED	264600
91013745 (1991, 1992)	Novel Protein sm. GBank g12650220(temb)CA0715791 - (Y10545) fused-cds5 [Eschrichia coli]	UNCLASSIFIED	264600
80503347 (1993, 1994)	Novel Protein sm. GBank g12650220(temb)CA0715791 - (Y10545) fused-cds5 [Eschrichia coli]	UNCLASSIFIED	264600
11397300 (1995, 1996)	Novel Protein sm. GBank g12650220(temb)CA0715791 - (Y10545) fused-cds5 [Eschrichia coli]	UNCLASSIFIED	264600

999	11765047 (1997, 1998)	Novel Protein sim. GBank g 25665697 p P4616 SVFCA_1HAEN - HYPOHETICAL PROTEIN H10150	UNCLASSIFIED	264682
1000	20727144 (1998, 2000)		UNCLASSIFIED	264602
1001	86673131 (2001, 2002)	Novel Protein sim. GBank g 22246994 p BAA20333 - (AB023377) KIAA0379 [Homo sapiens]	Contains protein domain (PF00023) - kinase	50432049, 264607, 264609, 264611, 264603, 264683, 264684, 264687, 264689, 264697, 21906789, 264692, 18108385, 22279000
1002	80189603 (2003, 2004)	Novel Protein sim. GBank g 586121 p P3709 TRHY_FABIT - TRICHOYALIN	silica	265009, 264369, 264620
1003	17933491 (2005, 2006)			265019
1004	16314867 (2007, 2008)	Novel Protein sim. GBank g 854095 p p CAA58337 - (K3413) U88 [Human herpesvirus 8]		264635
1005	79617144 (2009, 2010)	Novel Protein sim. GBank g 114073 p P07872 APT_ECQ1 - ADENINE PHOSPHORIBOSYLTRANSFERASE (APRT)	UNCLASSIFIED	264508
1006	37815426 (2011, 2012)	Novel Protein sim. GBank g 4062979 p g BAA36210.1 - (AB017139) sea urchin adult or maternal decarboxylase (Pseudomonas putida)	UNCLASSIFIED synthase	264259 264905
1007	79620877 (2013, 2014)			
1008	88094444 (2015, 2016)	Novel Protein sim. GBank g 2608007 p p CAA04607.1 - (AJ012069) putative trehalase synthase [Streptomyces celicofus]	synthase	265007, 264602, 264605, 264760, 264636
1009	57451289 (2017, 2018)	Novel Protein sim. GBank g 36330077 p p AF030113 - AMPA receptor binding protein [Rattus norvegicus]	Contains protein domain (PF00595) - kinase PDZ domain (Also known as DHR or GLGF)	264102, 264288
1010	84672537 (2019, 2020)	Novel Protein sim. GBank g 37468332 p p AF016307 - possible NADH-dependent oxidase, may function as a demethylase [Sinchizobium mellosi]	dehydrogenase	264592
1011	85548916 (2021, 2022)	Novel Protein sim. GBank g 2342647 p p AAB89581.1 - (U90653) DHHC-domain-containing cysteine-rich protein [Homo sapiens]	UNCLASSIFIED	33686032, 264605, 264764, 264768, 33695917, 264629
1012	95294446 (2023, 2024)	Novel Protein sim. GBank g 341341 p p CAA20222 - (AL031231) guanosine pentaphosphate synthetase/ adenosine nucleoside nucleosyltransferase [Streptomyces coelicolor]	Contains protein domain (PF00013) - phosphatase AT domain	33696032, 264605, 264606, 264601, 264602, 264605, 264762, 264768, 264769, 264689
1013	86095772 (2025, 2026)			264591, 21602768
1014	86600828 (2027, 2028)		UNCLASSIFIED	20331624, 265019, 265020

1015	95418879 (2029, 2030)	Novel Protein sim. GBank g14156952 (AF063095) - SELL [Mus musculus]	Contains protein domain (PF00040) Fibronectin type II domain	struct	22278904, 22278955, 56994075, 22278996, 22278999, 264256, 26331825, 26331828, 264907, 56182435, 264510, 264591, 264593, 60433356, 264594, 55812038, 264758, 21006754, 33657094, 265010, 264600, 265017, 265018, 265019, 18108351, 21906765, 21906766, 21906767, 21906768, 35811957, 32653985, 18108351, 18108377, 35811957, 32653985, 18108351, 18108377, 264593, 264600, 264603, 22279000, 264606, 264603
1016	7955694 (2031, 2032)	Novel Protein sim. GBank g12069899 (P44701MCPD, ECOLI - ACYL CARRIER PROTEIN PHOSPHODIESTERASE (ACP PHOSPHODIESTERASE))		esterase	
1017	11066213 (2033, 2034)	Novel Protein sim. GBank g15103843 (g15A479259.1) - (AP000059) 802aa long hypothetical oligopeptide-binding protein oppA [Aeropyrum pernix]	Contains protein domain (PF00406) Bacterial extracellular solute-binding proteins, family 5	transport	264600
1018	80072430 (2035, 2036)	Novel Protein sim. GBank g14493973 (emb1CAB3032.1) - (ALQ4559) predicted using hexExon, MAL37.14 (PFC0925w), Hypothetical protein, len: 489 aa [Plasmodium falciparum]			22278956, 29148827, 264563
1019	11703907 (2037, 2038)				
1020	80234432 (2039, 2040)			UNCLASSIFIED	264658 264508, 264509, 264512, 264600, 264762, 264769, 264669, 18108370, 264636, 264638, 264769
1021	37036243 (2041, 2042)	Novel Protein sim. GBank g14433407 (g14A028459.1) (AF12779) lipase-like hypothetical enzyme TreY [Rhizobium leguminosarum bv. viciae]		synthase	
1022	8052627 (2043, 2044)	Novel Protein sim. GBank g1781230 (mb1CAB38271) - (ZS987) hypothetical protein Rv3137 [Mycobacterium tuberculosis]	Contains protein domain (PF00459) host cell monophosphatases family	phosphatase	35695032, 264508, 265008, 265009, 264769, 18108387, 264563
1023	11399341 (2045, 2046)	Novel Protein sim. GBank g10777495 (J92083) - calcium transporting ATPase [Pichia angustis]	Contains protein domain (PF01122) E1-E2 ATPase	ATPase-associated	264503
1024	80057129 (2047, 2048)			UNCLASSIFIED	52846842, 33657402, 33657023, 18108379, 55811576, 264631, 264556, 264557, 264559, 18108388, 264566
1025	79644200 (2049, 2050)	Novel Protein sim. GBank g12483045 (mb1CA030556) - (AL031371) putative transport system permease protein [Streptomyces coelicolor]		transport	264693
1026	80029246 (2051, 2052)	Novel Protein sim. GBank g11749221 (g11C0232) (UVRD, HAEN - DNA HELICASE II [Heliobacterium salinarum])		helicase	264602
1027	17686934 (2053, 2054)	Novel Protein sim. GBank g145728 (g14A0686.1) (g14A07A - angiotensin vasopressin receptor A1/A1A-like)		UNCLASSIFIED	265017

1028	20397628 (2005, 2056)	Novel Protein sim. GBank g1791409jemb(CAA18603) - (AL021184) aza [Mycobacterium tuberculosis]	Contains protein domain (PF00330) - Acetate kinase family (acetate hydrolase)	UNCLASSIFIED	264600
1029	94055090 (2057, 2058)			UNCLASSIFIED	264595
1030	86095343 (2058, 2060)			UNCLASSIFIED	264510, 264512, 265009, 265010, 265011, 264600, 264602, 264603, 264605, 264769, 18108372, 18108374, 264805, 264908, 264909, 264909, 264955, 264992, 264930, 264634, 264636
1031	95289117 (2061, 2062)			UNCLASSIFIED	264689
1032	94073275 (2063, 2064)	Novel Protein sim. GBank g1403892refNP_000445.1 JGALK - galactokinase 1		UNCLASSIFIED	35680692, 264906, 264510, 18108354, 264687, 264769, 264689, 60431602, 18108385, 264486
1033	86164618 (2065, 2066)	Novel Protein sim. GBank g14982990 (AE000682) - hypothetical protein [Aureococcus anophagefferens]		UNCLASSIFIED	264908
1034	76249307 (2067, 2068)	Novel Protein sim. GBank g1405685 (U00007) - methionyl- tRNA synthetase [Escherichia coli]		UNCLASSIFIED	264682
1035	70165335 (2069, 2070)			UNCLASSIFIED	66712502, 264601, 264766
1036	85804668 (2071, 2072)			UNCLASSIFIED	264935, 264936, 264939, 60433438, 265019
1037	87360558 (2073, 2074)			UNCLASSIFIED	264634, 264634
1038	20431015 (2076, 2076)	Novel Protein sim. GBank g14005967 (AF030027) - 24 kidney disease-associated protein (Homo sapiens)	Contains protein domain (PF01477) - PLATLH2 domain	UNCLASSIFIED	264092, 264093, 264094, 264683, 264689, 263987
1039	87360021 (2077, 2078)	[Equus herpesvirus 4]		UNCLASSIFIED	264595
1040	80026840 (2079, 2080)	Novel Protein sim. GBank g14352035 (U91022) - DNA topoisomerase I [Fervidobacterium islandicum]	Contains protein domain (PF01131) - Prokaryotic DNA topoisomerase	isomerase	264595
1041	10156682 (2081, 2082)	Novel Protein sim. GBank g14256535 (p16AACB218.1) - kinase [Pyrococcus horikoshii]		kinase	264907
1042	11084375 (2083, 2084)	Novel Protein sim. GBank g14056289jemb(CAA66953) - (X9C359) Atp protein [Desophosphaerium]			264605
1043	80057136 (2085, 2086)	Novel Protein sim. GBank g14181701jemb(CAA10125) - (U08321) mink [Streptomyces reticuli]	Contains protein domain (PF00055) - ABC transporter	transport	264965, 264567
1044	80025852 (2087, 2088)			UNCLASSIFIED	265006, 264602, 265017
1045	52415482 (2089, 2090)	Novel Protein sim. GBank g14056890jemb(CAB52063.1) - (AL109739) hypothetical protein [Streptomyces coelicolor]		UNCLASSIFIED	29331823, 264637
1046	11754662 (2091, 2092)	Novel Protein sim. GBank g1454065jemb(CAA58337) - (X63413) U88 [Human herpesvirus 8]		helicase	264686
1047	37036258 (2093, 2094)	Novel Protein sim. GBank g142104710jemb(CAA74535.1) - (AB019033) orfSA [Pseudomonas sp.]		UNCLASSIFIED	264769
1048	79186400 (2095, 2096)	Novel Protein sim. GBank g14141341jemb(CAA320278) - (AL012322) hypothetical protein SC10H5.07 [Streptomyces coelicolor]	Contains protein domain (PF00220) - Neurohypophyseal hormones, N- terminal Domain	UNCLASSIFIED	264687
1049	81755108 (2097, 2098)	Novel Protein sim. GBank g15051630jemb(AAD30326.1) AF07372.2 EH domain-binding mitotic phosphoprotein [Homo sapiens]		UNCLASSIFIED	264805, 264634
1050	79471521 (2099, 2100)			UNCLASSIFIED	264686

1051	80475471 (2101, 2102)				UNCLASSIFIED	18108374, 264769, 265010, 265011, 264601, 265009, 264804, 264805, 264836, 18108351, 264832
1052	82442862 (2103, 2104)	Novel Protein sim. GBank g3132375pp35136SERA, BACSU - D-3- PHOSPHOGLYCERATE DEHYDROGENASE (PGDH- AL034980) dJ1158612.1 (zinc finger, X-linked, duplicated A) (Homo sapiens)	Contains protein domain (PF00339) - D-isomer specific 2-hydroxyacid dehydrogenases	- dehydrogenase		264308, 264762, 264687, 264486
1053	94851640 (2105, 2106)	Novel Protein sim. GBank g1021676jg139embICAB4671.1 - (AL034980) dJ1158612.1 (zinc finger, X-linked, duplicated A) (Homo sapiens)				264686, 18108374, 26331824, 83373044, 21906754, 52645156, 5182435, 264689, 26331827, 27486281, 35698052, 21906765, 35698423, 21906768, 56182575, 21906769, 58811957, 87168518, 35692896, 22278897, 265020, 265011, 265021, 265022, 265007, 265012, 265008, 264806, 26482, 264906, 52644155, 264909, 264886, 264908
1054	75502025 (2107, 2108)	Novel Protein sim. GBank g10542508jg1A038584.1 (AF14560) - BcDNA GH02833 (Orcophila melanogaster)	Contains protein domain (PF00270) - DEAD/DEAH box helicase	helicase	UNCLASSIFIED	264686
1055	80394138 (2109, 2110)	Novel Protein sim. GBank g1021676jg139embICAB4671.1 - (AL034980) dJ1158612.1 (zinc finger, X-linked, duplicated A) (Homo sapiens)	Contains protein domain (PF00270) - DEAD/DEAH box helicase	helicase	UNCLASSIFIED	264907, 264602, 264681, 264285, 21908768, 33657105, 55810764, 35695855, 264631
1056	17682319 (2111, 2112)	Novel Protein sim. GBank g1021676jg139embICAB4671.1 - (AL034980) dJ1158612.1 (zinc finger, X-linked, duplicated A) (Homo sapiens)	(O88033) RNA polymerase sigma 70 factor (Pseudomonas fluorescens)	mapolymerase		264905
1057	85667216 (2113, 2114)	Novel Protein sim. GBank g1226281 (U53008) - No definition line found (Caenorhabditis elegans)			UNCLASSIFIED	264682
1058	80767576 (2115, 2116)	Novel Protein sim. GBank g1170018ppP4680jgREA, MYCLE - TRANSCRIPTION ELONGATION FACTOR GREY (TRANSCRIPT CLEAVAGE FACTORS GREY)				264764
1059	9462754 (2117, 2118)	Novel Protein sim. GBank g1226281 (U53008) - No definition line found (Caenorhabditis elegans)			transcript factor	35696032, 35695855, 265008, 264636
1060	79481169 (2119, 2120)	Novel Protein sim. GBank g1226281 (U53008) - No definition line found (Caenorhabditis elegans)			glycoprotein	25116499, 264681, 264653, 264687
1061	11034025 (2121, 2122)	Novel Protein sim. GBank g1021676jg139embICAB4671.1 - (AL034980) dJ1158612.1 (zinc finger, X-linked, duplicated A) (Homo sapiens)			phosphatase	264634
1062	38267937 (2123, 2124)	Novel Protein sim. GBank g13334200jg104684jSCSP, SOLTU - GLYCINE DEHYDROGENASE (DECARBOXYLATING) PRECURSOR (GLYCINE DECARBOXYLASE) (GLYCINE CLEAVAGE SYSTEM P-PROTEIN)			dehydrogenase	264583
1063	9460451 (2125, 2126)	Novel Protein sim. GBank g1226281 (U53008) - No definition line found (Caenorhabditis elegans)				264908
1064	76591783 (2127, 2128)	Novel Protein sim. GBank g1021676jg139embICAB4671.1 - (AL034980) dJ1158612.1 (zinc finger, X-linked, duplicated A) (Homo sapiens)				265007, 265008, 18108351, 18108385

1065	50031208 (2129, 2130)	Novel Protein sim. GBank gij2120959[pe] (ST0882 - <i>Staphylococcus aureus</i>)	transferase	264600, 264602, 264680
1066	17866679 (2131, 2132)	Novel Protein sim. GBank gij2506362[pe]P156042[DNLI] ECOLI - DNA LIGASE (POLYDEOXYRIBONUCLEOTIDE SYNTHASE (NAD+))	synthase	265009
1067	10132178 (2133, 2134)	Novel Protein sim. GBank gij4007669[emb]CAA22355[- (ALD34443) putative oxidoreductase (Streptomyces coelicolor)	reductase	264909
1068	82062057 (2135, 2136)	Novel Protein sim. GBank gij438584[dm]BAAT6770.1[- (AB023143) KIAA0928 protein (Homo sapiens)	UNCLASSIFIED	264658, 18108362, 264455, 264600, 264760
1069	83002854 (2137, 2138)	Novel Protein sim. GBank gij1750397 (J01281) - glutamate synthase large subunit (<i>Pseudomonas aeruginosa</i>)	UNCLASSIFIED	264604
1070	82101992 (2139, 2140)	Novel Protein sim. GBank gij1200346[pe]P15932[FLG, SALT - FLAGELLAR HOOK-ASSOCIATED PROTEIN 1 (HAP1)]	UNCLASSIFIED	264604, 264760
1071	20710589 (2141, 2142)	Novel Protein sim. GBank gij1750397 (J01281) - glutamate synthase large subunit (<i>Pseudomonas aeruginosa</i>)	synthase	264602
1072	82356540 (2143, 2144)	Novel Protein sim. GBank gij477532[pe]J049175 - MotC B protein - mouse (fragment)	UNCLASSIFIED	264687, 264688, 21006764, 35608052, 35698917, 35695555, 264600, 264601, 264602, 265000, 264605, 264508, 264905, 264600, 264606, 264762, 264528, 264768
1073	73814400 (2145, 2146)	Novel Protein sim. GBank gij477532[pe]J049175 - MotC B protein - mouse (fragment)	synthase	264906
1074	80105892 (2147, 2148)	Novel Protein sim. GBank gij3383109[emb]CAA7694[- (V17820) CALO protein (<i>Drosophila melanogaster</i>)	UNCLASSIFIED	56994075, 22276998, 264594, 264757, 264596, 265016, 265019, 264681, 264389, 264688, 265020, 18108364, 18108374
1075	81850293 (2149, 2150)	Novel Protein sim. GBank gij477532[pe]J049175 - MotC B protein - mouse (fragment)	ATPase-associated	264769
1076	80477284 (2151, 2152)	Novel Protein sim. GBank gij1176203[pe]P46442[HYCM, ECOLI - HYPOTHETICAL 43.1 KD PROTEIN IN RPLM-HHOA INTERGENIC REGION (P375)	kinase	264905
1077	79831334 (2153, 2154)	Novel Protein sim. GBank gij403487[pe]P44472[TUOM, AGRI - PUTATIVE HYDROXYPYRUVATE REDUCTASE	UNCLASSIFIED	264600
1078	20286974 (2155, 2156)	Novel Protein sim. GBank gij51343[emb]CAA2028[- (ALD31760) hypothetical protein SC0410.09 (<i>Streptomyces coelicolor</i>)	UNCLASSIFIED	18108394, 264769, 264834, 264636
1079	80494718 (2157, 2158)	Novel Protein sim. GBank gij51343[emb]CAA2028[- (ALD31760) hypothetical protein SC0410.09 (<i>Streptomyces coelicolor</i>)	UNCLASSIFIED	264684
1080	11767168 (2159, 2160)	Novel Protein sim. GBank gij51343[emb]CAA2028[- (ALD31760) hypothetical protein SC0410.09 (<i>Streptomyces coelicolor</i>)	UNCLASSIFIED	83373044, 265019, 22276002, 264482, 18108351, 264652, 264908, 264693, 264487, 264758, 264768, 264769, 21906767, 264511, 264910, 264634, 264635, 264905, 264636, 264906, 264637, 264907, 264906, 264764, 264638, 20281099, 264706, 264995
1081	94747080 (2161, 2162)	Novel Protein sim. GBank gij51343[emb]CAA2028[- (ALD31760) hypothetical protein SC0410.09 (<i>Streptomyces coelicolor</i>)	UNCLASSIFIED	
1082	81490656 (2163, 2164)	Novel Protein sim. GBank gij51343[emb]CAA2028[- (ALD31760) hypothetical protein SC0410.09 (<i>Streptomyces coelicolor</i>)	UNCLASSIFIED	

1083	187445717 (2165, 2168)	Novel Protein sim, GBank g117229459p105210YIN_MJCTUC - HYPOTHETICAL 44.6 KD PROTEIN C7427.23		UNCLASSIFIED	60424179, 264905, 264600, 264510, 60432229, 264759, 17186474, 264605, 264789, 264689, 18108364, 18108378, 35695935, 264636 264169
1084	37799306 (2167, 2168)	Novel Protein sim, GBank g9418349p12057MVAL_ECOLI - PUTATIVE COLANIC ACID BIOSYNTHESIS GLYCOSYL TRANSFERASE VCAI		UNCLASSIFIED	
1085	86475368 (2169, 2170)	Novel Protein sim, GBank g11990190 (J09204) - heat shock protein 60 [Tetrahymena thermophila]	Contains protein domain (PF00118) - eph		60432229, 264687
1086	79603259 (2171, 2172)	Novel Protein sim, GBank g11729585p146176RL14_ACTYK - 5.5 RIBOSOMAL PROTEIN L14	Contains protein domain (PF00238) - ribosomal protein L14		264486
1087	79603979 (2173, 2174)	Novel Protein sim, GBank g14160198p16CXA154311 - (AL008883) d322716.3 (novel CHROMOX family protein) [Bromo sapientis]	Contains protein domain (PF00385) - helicase chromo (CHROMATIN Organization Modifier) domain		26331827, 264693
1088	79645963 (2175, 2176)	Novel Protein sim, GBank g12983155 (AE000693) - phosphoglucosyltransferase [Streptomyces coelicolor]	UNCLASSIFIED		264905, 264601, 18108387
1089	80216800 (2177, 2178)	Novel Protein sim, GBank g14981758p14039260_14500177 - (AE001776) NACH dehydrogenase, 30 kDa subunit, putative [Thermotoga maritima]	Contains protein domain (PF03329) - Respiratory-chain NADH dehydrogenase, 30 kDa subunit	UNCLASSIFIED	264468, 264511, 265011, 264882, 264768, 264689, 121900764, 35695917, 265020, 32833966, 18106370, 35695935
1090	11083625 (2179, 2180)	Novel Protein sim, GBank g1400768p16CXA22666 - (AL034443) putative oxidoreductase [Streptomyces coelicolor]			264604
1091	72917471 (2181, 2182)	Novel Protein sim, GBank g12495652p1772381YLCD_ECOLI - HYPOTHETICAL 44.3 KD PROTEIN IN NFRB-PHEP INTERGENIC REGION PRECURSOR	UNCLASSIFIED		264637
1092	80252768 (2183, 2184)	Novel Protein sim, GBank g12960098p16CXA17996.11 - (AL022121) nh [Myobacterium tuberculosis]	Contains protein domain (PF00730) - Endonuclease III		264566
1093	80496304 (2185, 2186)	Novel Protein sim, GBank g11001642p16CXA103731 - (J04002) d31P-tyrosine phosphatase [Synchrotrons sp.]	UNCLASSIFIED		264789, 35695917, 35695935, 264600, 264602, 264603, 18108351
1094	10860972 (2187, 2188)	Novel Protein sim, GBank g1468358p16CXA40855.11 - (AL04829) putative aspartate glycosylase [Streptomyces coelicolor]	Contains protein domain (PF04453) - Bacterial regulatory proteins, deoR family		264696
1095	97457250 (2189, 2190)	Novel Protein sim, GBank g1468358p16CXA40855.11 - (AL04829) putative aspartate glycosylase [Streptomyces coelicolor]			264906, 265007, 264588, 264600, 264602, 264603, 264604, 264605, 264762, 264766, 264789, 264636, 264536, 18108387, 264600, 264602, 264603, 264604, 264605, 35695917, 264692, 264631
1096	80025977 (2191, 2192)	Novel Protein sim, GBank g11150011p19209BIBO_BACSH - BIOTIN SYNTHASE (BIOTIN SYNTHETASE)	synthase		265019
1097	76239550 (2193, 2194)	Novel Protein sim, GBank g1114135p19209BIBO_BACSH - BIOTIN SYNTHASE	UNCLASSIFIED		264687
1098	75186424 (2195, 2196)	Novel Protein sim, GBank g1114135p19209BIBO_BACSH - BIOTIN SYNTHASE	synthase		

1099	35023533 (2197, 2198)	Novel Protein sim. G.Bank g 39151449 Q3017 TRMD_MYCLE - TRNA (GLUTAMINE- N1)-METHYLTRANSFERASE (M1G- METHYLTRANSFERASE) (TRNA [G307] METHYLTRANSFERASE)	UNCLASSIFIED	264603
1100	85736571 (2196, 2200)	Novel Protein sim. G.Bank g 3023255 PQ4420 ACOD_MESAU - ACYL-COA DESATURASE (STEAROYL-COA DESATURASE) (FATTY ACID DESATURASE) (DELTA9-DESATURASE)	desaturase	264259, 264636
1101	80491857 (2201, 2202)	Novel Protein sim. G.Bank g 11747339 p43012 TOP1_LHAEN - DNA TOPISOMERASE (OMEGA-PROTEIN) (RELAXING ENZYME) (UNWINDING ENZYME) (SWIVELASE)	Topoisomerase DNA binding C4 zinc finger	264769
1102	79777814 (2203, 2204)	Novel Protein sim. G.Bank g 11747339 p43012 TOP1_LHAEN - DNA TOPISOMERASE (OMEGA-PROTEIN) (RELAXING ENZYME) (UNWINDING ENZYME) (SWIVELASE)	isomerase	264769
1103	81697255 (2205, 2206)	Novel Protein sim. G.Bank g 11747339 p43012 TOP1_LHAEN - DNA TOPISOMERASE (OMEGA-PROTEIN) (RELAXING ENZYME) (UNWINDING ENZYME) (SWIVELASE)	isomerase	264769
1104	35003115 (2207, 2208)	Novel Protein sim. G.Bank g 11747339 p43012 TOP1_LHAEN - DNA TOPISOMERASE (OMEGA-PROTEIN) (RELAXING ENZYME) (UNWINDING ENZYME) (SWIVELASE)	isomerase	264769
1105	80255121 (2209, 2210)	Novel Protein sim. G.Bank g 11747339 p43012 TOP1_LHAEN - DNA TOPISOMERASE (OMEGA-PROTEIN) (RELAXING ENZYME) (UNWINDING ENZYME) (SWIVELASE)	isomerase	264769
1106	78514110 (2211, 2212)	Novel Protein sim. G.Bank g 11747339 p43012 TOP1_LHAEN - DNA TOPISOMERASE (OMEGA-PROTEIN) (RELAXING ENZYME) (UNWINDING ENZYME) (SWIVELASE)	isomerase	264769
1107	80470019 (2213, 2214)	Novel Protein sim. G.Bank g 11747339 p43012 TOP1_LHAEN - DNA TOPISOMERASE (OMEGA-PROTEIN) (RELAXING ENZYME) (UNWINDING ENZYME) (SWIVELASE)	isomerase	264769
1108	80440616 (2215, 2216)	Novel Protein sim. G.Bank g 11747339 p43012 TOP1_LHAEN - DNA TOPISOMERASE (OMEGA-PROTEIN) (RELAXING ENZYME) (UNWINDING ENZYME) (SWIVELASE)	isomerase	264769
1109	80064615 (2217, 2218)	Novel Protein sim. G.Bank g 11747339 p43012 TOP1_LHAEN - DNA TOPISOMERASE (OMEGA-PROTEIN) (RELAXING ENZYME) (UNWINDING ENZYME) (SWIVELASE)	isomerase	264769
1110	80503354 (2219, 2220)	Novel Protein sim. G.Bank g 11747339 p43012 TOP1_LHAEN - DNA TOPISOMERASE (OMEGA-PROTEIN) (RELAXING ENZYME) (UNWINDING ENZYME) (SWIVELASE)	isomerase	264769
1111	80071744 (2221, 2222)	Novel Protein sim. G.Bank g 11747339 p43012 TOP1_LHAEN - DNA TOPISOMERASE (OMEGA-PROTEIN) (RELAXING ENZYME) (UNWINDING ENZYME) (SWIVELASE)	isomerase	264769
1112	80071744 (2221, 2222)	Novel Protein sim. G.Bank g 11747339 p43012 TOP1_LHAEN - DNA TOPISOMERASE (OMEGA-PROTEIN) (RELAXING ENZYME) (UNWINDING ENZYME) (SWIVELASE)	isomerase	264769
1113	82465352 (2225, 2226)	Novel Protein sim. G.Bank g 11747339 p43012 TOP1_LHAEN - DNA TOPISOMERASE (OMEGA-PROTEIN) (RELAXING ENZYME) (UNWINDING ENZYME) (SWIVELASE)	isomerase	264769
1114	14898014 (2227, 2228)	Novel Protein sim. G.Bank g 11747339 p43012 TOP1_LHAEN - DNA TOPISOMERASE (OMEGA-PROTEIN) (RELAXING ENZYME) (UNWINDING ENZYME) (SWIVELASE)	isomerase	264769
1115	11765383 (2229, 2230)	Novel Protein sim. G.Bank g 11747339 p43012 TOP1_LHAEN - DNA TOPISOMERASE (OMEGA-PROTEIN) (RELAXING ENZYME) (UNWINDING ENZYME) (SWIVELASE)	isomerase	264769
1116	17841132 (2231, 2232)	Novel Protein sim. G.Bank g 11747339 p43012 TOP1_LHAEN - DNA TOPISOMERASE (OMEGA-PROTEIN) (RELAXING ENZYME) (UNWINDING ENZYME) (SWIVELASE)	isomerase	264769

1117	95305465 (2233, 2234)	Novel Protein sim. GBank g[13255/54]mim:CA34089g - (270200) US antiRNP specific 200KD protein [Homo sapiens]	Contains protein domain (PF00770) - DEAD/DEAF box helicase	18103396, 264485, 265994, 264485, 2596266, 26594075, 2277895, 2227896, 2277897, 2654293, 26331822, 26162181, 26331824, 26331827, 26331833, 26330532, 60432260, 26331827, 26331833, 26330532, 33656970, 26146496, 264508, 264905, 264909, 264906, 264007, 264908, 68712502, 264909, 52844045, 56182435, 264510, 264511, 265005, 265007, 264512, 265008, 265009, 264910, 60170831, 264591, 264592, 264593, 60433356, 264594, 60433436, 264595, 55812038, 264596, 264758, 264759, 21906794, 33106954, 33657084, 265011, 87168359, 264600, 264601, 264602, 264603, 265017, 264604, 265018, 264605, 265019, 264760, 55811150, 264681, 264762, 18103391, 264448, 264682, 264764, 264683, 264286, 264389, 264684, 264687, 264766, 264767, 264685, 264687, 264768, 264769, 264688, 1810953, 264689, 264690, 21906766, 21906767, 21906768, 21906769, 55811957, 26148784, 33669917, 265020, 265021, 265022, 60170615, 264690, 264691, 33657023, 264682, 264693, 65274620, 33657109, 33657162, 27488281, 27488282, 27488284, 33657249, 27488255, 35695763, 264691
1118	7953326 (2235, 2236)		UNCLASSIFIED	264691
1119	7942463 (2237, 2238)		UNCLASSIFIED	264697
1120	7948045 (2239, 2240)	Novel Protein sim. GBank g[1420387]mim:CA846679.11 - (A1243459) proteophosphoglycan [Leishmania major]	collagen	29331827, 265018, 265019, 264687, 265021, 60170615, 18109387
1121	7947176 (2241, 2242)	Novel Protein sim. GBank g[1644450] (U67864) - MEK-3	Contains protein domain (PF00013) - K4 domain	264683, 264632, 18108388
1122	79455246 (2243, 2244)	Carionabidula elegans		
1123	79537118 (2245, 2246)	Novel Protein sim. GBank g[188800]mim:J157768 - 3, dehydroquinolate synthase (EC 4.6.1.3) - Mycobacterium tuberculosis	UNCLASSIFIED	264693, 27486265
1124	79811596 (2247, 2248)		UNCLASSIFIED	264699
1125	79757861 (2249, 2250)		UNCLASSIFIED	264610
1126	79758914 (2251, 2252)	Novel Protein sim. GBank g[138154]mim:PP0364VGG, BPPHX - MAJOR SPIKE PROTEIN (G PROTEIN) (GPG)	eph	264905, 264908, 264910
1127	11609530 (2253, 2254)	Novel Protein sim. GBank g[100274]mim:CA844358.11 - (A1242930) DNA polymerase I [Methylobacterium sp. DM4]	UNCLASSIFIED	264682
1128	8304885 (2255, 2256)		UNCLASSIFIED	264511

1167	79663662 (2333, 2334)	Novel Protein sim. GBank g12560433(bj)BAJ231381 - (D78414) ppGpp hydrolase (Slaptp/ococcus aureus)		kinase	264458
1168	80094678 (2335, 2336)			UNCLASSIFIED	264253, 26331827, 56182435, 60433438, 265016, 33657023, 35695855, 264566
1169	11805403 (2337, 2338)			UNCLASSIFIED	264681
1170	21632244 (2339, 2340)			UNCLASSIFIED	264602
1171	26434582 (2341, 2342)			UNCLASSIFIED	264556
1172	79610113 (2343, 2344)	Novel Protein sim. GBank g1272014 (AF028249) - precollagen D [Mytilus edulis]		UNCLASSIFIED	55810764, 35666002, 264634, 264486
1173	80235713 (2345, 2346)	Novel Protein sim. GBank g1254653(bj)BAJ229401 - (AB007352) Bm. trophoblast [Bm. trophoblast]			264508, 264306, 264307, 264509, 264591, 264632, 264638, 264639
1174	20252077 (2347, 2348)	Novel Protein sim. GBank g12511027(m)CAA175201 - (AL021958) mmsA [Mycobacterium tuberculosis]		dehydrogenase	264600
1175	20711847 (2349, 2350)	Novel Protein sim. GBank g118333ppP2324(DCP, ENTCL - INDOLE-3-PYRUVATE DECARBOXYLASE (INDOLEPYRUVATE DECARBOXYLASE)		carboxylase	264601
1176	80252845 (2351, 2352)	Novel Protein sim. GBank g11145220 (U34956) - phosphoribosylformylglycinamide synthase [Mycobacterium tuberculosis]		synthase	264508, 264505, 264593, 264602, 264605
1177	80064647 (2353, 2354)	Novel Protein sim. GBank g118791ppP28643PABG, CUPJA - 3-OXOACYL-ACYL-CARRIER PROTEIN REDUCTASE PRECURSOR (3-KETOACYL-ACYL CARRIER PROTEIN REDUCTASE)		reductase	264605
1178	84128641 (2355, 2356)	Novel Protein sim. GBank g1255779ppP-003594.1(gf)FC1 - familial intrahepatic cholestasis 1, (progressive, liver disease and benign recurrent)			65274872, 18103388, 22278958, 22278959, 23031626, 204508, 264608, 264826, 33657402, 33109954, 264765, 21906765, 21906768, 21906768, 55811957, 33657023, 264628, 55811576, 35696423, 264636, 264656, 55812323, 60432113, 22279000, 22279002, 264603
1179	80055575 (2357, 2358)	Novel Protein sim. GBank g12600009ppP(CAA17988 - 1) - (AL021212) oppA [Mycobacterium tuberculosis]		transport	
1180	11794446 (2359, 2360)	Novel Protein sim. GBank g12558613ppP(CAA041871) - (A001483) dehydroquinase dehydroalase (Streptomyces coelicolor)		synthase	264638
1181	17940363 (2361, 2362)	Novel Protein sim. GBank g15420387ppP(CAB46673 - 1) - (A1243459) proteophosphoglycan [Leishmania major]		UNCLASSIFIED	265017
1182	81943254 (2363, 2364)				265007, 265009, 264564, 264309, 264693
1183	79571041 (2365, 2366)	Novel Protein sim. GBank g14091877 (AF061331) - alpha helicoidase (Schistosoma japonicum)		UNCLASSIFIED	264686, 35696423, 264638, 18108385
1184	82555933 (2367, 2368)	Novel Protein sim. GBank g11212121ppP(g11212121) - chitinase (EC 3.2.1.14) precursor - beet			264602
1185	79491182 (2369, 2370)			glycoprotein	265367

1209	79185742 (2417, 2418)	Novel Protein sim. GBank g11175033spP44398XYLA_HAEIN - XYLOSE ISOMERASE	Contains protein domain (PF00259) - Xylose isomerase	isomerase	264687, 264688
1210	9626884 (2418, 2420)				
1211	9465655 (2421, 2422)	Novel Protein sim. GBank g1021098ipJ150388 - hypothetical protein c246 - <i>Escherichia coli</i>		UNCLASSIFIED transferase	264697, 264693 264691, 264692, 264695
1212	79167929 (2423, 2424)	Novel Protein sim. GBank g1080629mbiCAB07859j - (293785) predicted using GeneFinder; similar to RNA recognition motif (aka RRM, RBD, or RNP domain); cDNA EST EMBL101662 comes from this gene; cDNA EST EMBL X75825 comes from this gene; cDNA EST EMBL D27559 comes from this gene	Contains protein domain (PF01412) - Putative GTP-ase activating protein for Arf		264689, 263967
1213	79096633 (2425, 2426)	Novel Protein sim. GBank g1026262ipJ1505375A - <i>vir</i>		kinase	264699
1214	10144306 (2427, 2428)	Novel Protein sim. GBank g1026262ipJ1505375A - <i>vir</i>		UNCLASSIFIED	264698
1215	80050106 (2429, 2430)	Novel Protein sim. GBank g1026262ipJ1505375A - <i>vir</i>		UNCLASSIFIED	265009, 264601, 264602, 264603, 35657169
1216	20435324 (2431, 2432)	Novel Protein sim. GBank g1026262ipJ1505375A - <i>vir</i>		transferase	264604
1217	85011344 (2433, 2434)				
1218	11093600 (2435, 2436)	Novel Protein sim. GBank g1026262ipJ1505375A - <i>vir</i>		UNCLASSIFIED	264699, 264698, 264697, 264696, 264691, 264693, 264692, 264695, 264694
1219	91216325 (2437, 2438)	Novel Protein sim. GBank g1026262ipJ1505375A - <i>vir</i>		dehydrogenase	264601
1220	91241524 (2439, 2440)	Novel Protein sim. GBank g102403150ipJAX4936 1) - (AB020720) KIA0913 protein [Homo sapiens]		oncogene	56181698, 20331182, 60432289, 264601, 264692, 264629
1221	83045055 (2441, 2442)	Novel Protein sim. GBank g102403150ipJAX4936 1) - (AB020720) KIA0913 protein [Homo sapiens]		dehydrogenase	56181698, 20331182, 60432289, 264601, 264692, 264629
1222	20711866 (2443, 2444)	Novel Protein sim. GBank g102403150ipJAX4936 1) - (AB020720) KIA0913 protein [Homo sapiens]	Contains protein domain (PF00486) - Transcriptional regulatory protein, C terminal	phosphatase	264601
1223	11019647 (2445, 2446)	Novel Protein sim. GBank g102403150ipJAX4936 1) - (AB020720) KIA0913 protein [Homo sapiens]			
1224	80432645 (2447, 2448)	Novel Protein sim. GBank g102403150ipJAX4936 1) - (AB020720) KIA0913 protein [Homo sapiens]	Contains protein domain (PF01472) - PJA domain	kinase	264693, 264690, 264601, 264603, 264605, 264698, 10100375, 264635, 10100387

1226	80234427 (2445, 2450)	Novel Protein sim. GBank g1205050(emb CA00833 - (2554-35) hypothetical protein Rv044c Mycobacterium tuberculosis)			264768 264905, 264512, 264689
1228	80237516 (2451, 2452)	Novel Protein sim. GBank g1700768(p P133 FBN1_BOVIN - FIBRILLIN 1 PRECURSOR (NP240))	Contains protein domain (PF00073) - AcBlaC-DNAse family		264608, 264637, 264639
1229	79422138 (2453, 2454)	Novel Protein sim. GBank g1163360(p D JAA11811) - Novel Protein sim. GBank g1163360(p D JAA11811) - (D90917) acriflavine resistance protein (Synechococcus sp.)	Contains protein domain (PF00018) - TCP-1cp180 chaperonin family		264605, 264634
1228	79206027 (2455, 2456)	Novel Protein sim. GBank g1163360(p D JAA11811) - (D90917) acriflavine resistance protein (Synechococcus sp.)	Contains protein domain (PF00018) - TCP-1cp180 chaperonin family		264605, 264634
1229	94329135 (2457, 2458)	Novel Protein sim. GBank g1163360(p D JAA11811) - (D90917) acriflavine resistance protein (Synechococcus sp.)	Contains protein domain (PF00018) - TCP-1cp180 chaperonin family		264605, 264634
1230	80046357 (2459, 2460)	Novel Protein sim. GBank g1163360(p D JAA11811) - (D90917) acriflavine resistance protein (Synechococcus sp.)	Contains protein domain (PF00018) - TCP-1cp180 chaperonin family		264605, 264634
1231	79643141 (2461, 2462)	Novel Protein sim. GBank g1163360(p D JAA11811) - (D90917) acriflavine resistance protein (Synechococcus sp.)	Contains protein domain (PF00018) - TCP-1cp180 chaperonin family		264605, 264634
1232	79653104 (2463, 2464)	Novel Protein sim. GBank g1163360(p D JAA11811) - (D90917) acriflavine resistance protein (Synechococcus sp.)	Contains protein domain (PF00018) - TCP-1cp180 chaperonin family		264605, 264634
1233	80255179 (2465, 2466)	Novel Protein sim. GBank g1163360(p D JAA11811) - (D90917) acriflavine resistance protein (Synechococcus sp.)	Contains protein domain (PF00018) - TCP-1cp180 chaperonin family		264605, 264634
1234	79242158 (2467, 2468)	Novel Protein sim. GBank g1163360(p D JAA11811) - (D90917) acriflavine resistance protein (Synechococcus sp.)	Contains protein domain (PF00018) - TCP-1cp180 chaperonin family		264605, 264634
1235	79914423 (2469, 2470)	Novel Protein sim. GBank g1163360(p D JAA11811) - (D90917) acriflavine resistance protein (Synechococcus sp.)	Contains protein domain (PF00018) - TCP-1cp180 chaperonin family		264605, 264634
1236	81927147 (2471, 2472)	Novel Protein sim. GBank g1163360(p D JAA11811) - (D90917) acriflavine resistance protein (Synechococcus sp.)	Contains protein domain (PF00018) - TCP-1cp180 chaperonin family		264605, 264634
1237	83371782 (2473, 2474)	Novel Protein sim. GBank g1163360(p D JAA11811) - (D90917) acriflavine resistance protein (Synechococcus sp.)	Contains protein domain (PF00018) - TCP-1cp180 chaperonin family		264605, 264634
1238	81411577 (2475, 2476)	Novel Protein sim. GBank g1163360(p D JAA11811) - (D90917) acriflavine resistance protein (Synechococcus sp.)	Contains protein domain (PF00018) - TCP-1cp180 chaperonin family		264605, 264634
1239	82197449 (2477, 2478)	Novel Protein sim. GBank g1163360(p D JAA11811) - (D90917) acriflavine resistance protein (Synechococcus sp.)	Contains protein domain (PF00018) - TCP-1cp180 chaperonin family		264605, 264634
1240	80497259 (2479, 2480)	Novel Protein sim. GBank g1163360(p D JAA11811) - (D90917) acriflavine resistance protein (Synechococcus sp.)	Contains protein domain (PF00018) - TCP-1cp180 chaperonin family		264605, 264634
1241	80020711 (2481, 2482)	Novel Protein sim. GBank g1163360(p D JAA11811) - (D90917) acriflavine resistance protein (Synechococcus sp.)	Contains protein domain (PF00018) - TCP-1cp180 chaperonin family		264605, 264634
1242	79775890 (2483, 2484)	Novel Protein sim. GBank g1163360(p D JAA11811) - (D90917) acriflavine resistance protein (Synechococcus sp.)	Contains protein domain (PF00018) - TCP-1cp180 chaperonin family		264605, 264634

1243	70776458 (2485, 2486)	Novel protein sm. GBank gll356975(embC:AA11927.1) - (AU031124) branched-chain amino acid aminotransferase [Streptomyces coelicolor]	UNCLASSIFIED	18108374, 26565917, 26565955, 26566009, 264408, 264609
1244	10284242 (2487, 2488)	Novel Protein sm. GBank gll2970646 (AF051945) - Xln [Mus musculus]	UNCLASSIFIED	264681
1245	80437103 (2485, 2490)	Novel Protein sm. GBank gll456633(embC:AA76357.1) - (AB016787) cytochrome c ubiquinol oxidase B	oxidase	264768
1246	80059321 (2491, 2492)	Novel Protein sm. GBank gll358184(embC:AA20805) - (AU031541) putative phenylalanine-HRNA synthetase beta chain [Streptomyces coelicolor]	UNCLASSIFIED	264604, 264636, 264657, 264664
1247	80064431 (2493, 2494)	Novel Protein sm. GBank gll562184 (AE000842) - adhesion protein [Methanobacterium thermoautotrophicum]		264756, 264605, 264639
1248	88070353 (2485, 2486)	Novel Protein sm. GBank gll152403(embC:AA76357.1) - (AF161317) FRUCTOSE-1,6-BISPHOSPHATE 1-PHOSPHOHYDROLYASE (FBPASE)	UNCLASSIFIED	18108332, 264259, 26331626, 264106, 264508, 264907, 264428, 265005, 60433366, 264757, 264758, 2189754, 265010, 265011, 265018, 265019, 264720, 18108351, 18108352, 18108353, 18108354, 18108355, 18108356, 264830, 18108376, 18108377, 264830, 265010, 264660, 264603, 264691, 18108376
1249	80056657 (2497, 2498)	Novel Protein sm. GBank gll2791407(embC:AA16001) - (AU021184) hypothetical protein Rv1472 [Mycobacterium tuberculosis]	transport	264689
1250	12684385 (2496, 2500)	Novel Protein sm. GBank gll112785(embC:AA16001) - (AF161317) NRAMP METHYLADENINE GLYCOSIDASE [3-METHYLADENINE-DNA GLYCOSYLASE I, CONSTITUTIVE] (TAG II)	UNCLASSIFIED	264683, 263976
1251	79859448 (2501, 2502)	Novel Protein sm. GBank gll567076(embC:AA16001) - (AF161317) NRAMP manganese transport protein MntA [Salmonella typhimurium]	glycoprotein	264609
1252	79450807 (2503, 2504)	Novel Protein sm. GBank gll103169(embC:AA16001) - (AF161317) protein unknep1 - fruit fly (Drosophila melanogaster)	UNCLASSIFIED	264683, 263976
1253	80050121 (2505, 2506)	Novel Protein sm. GBank gll103169(embC:AA16001) - (AF161317) protein unknep1 - fruit fly (Drosophila melanogaster)	UNCLASSIFIED	264683, 263976
1254	80716767 (2507, 2508)	Novel Protein sm. GBank gll103169(embC:AA16001) - (AF161317) protein unknep1 - fruit fly (Drosophila melanogaster)	UNCLASSIFIED	264683, 263976
1255	79169728 (2509, 2510)	Novel Protein sm. GBank gll103169(embC:AA16001) - (AF161317) protein unknep1 - fruit fly (Drosophila melanogaster)	UNCLASSIFIED	264683, 263976
1256	87389508 (2511, 2512)	Novel Protein sm. GBank gll103169(embC:AA16001) - (AF161317) protein unknep1 - fruit fly (Drosophila melanogaster)	UNCLASSIFIED	264683, 263976
1257	80201435 (2513, 2514)	Novel Protein sm. GBank gll103169(embC:AA16001) - (AF161317) protein unknep1 - fruit fly (Drosophila melanogaster)	UNCLASSIFIED	264683, 263976
1258	20708150 (2515, 2516)	Novel Protein sm. GBank gll103169(embC:AA16001) - (AF161317) protein unknep1 - fruit fly (Drosophila melanogaster)	UNCLASSIFIED	264683, 263976
1259	80186012 (2517, 2518)	Novel Protein sm. GBank gll103169(embC:AA16001) - (AF161317) protein unknep1 - fruit fly (Drosophila melanogaster)	UNCLASSIFIED	264683, 263976
1260	80084066 (2519, 2520)	Novel Protein sm. GBank gll103169(embC:AA16001) - (AF161317) protein unknep1 - fruit fly (Drosophila melanogaster)	UNCLASSIFIED	264683, 263976

1316	95361609 (2631, 2632)	Novel Protein sim. GBank g16580407/dbjBAA02887.1] - (AB026858) KIAA1035 protein [Homo sapiens]	kinase	56182515, 56181586, 20381171, 29331822, 29331822, 80424369, 80432229, 2650918, 52840045, 264691, 80432229, 2650918, 265018, 55911150, 56115162, 21906765, 21908787, 21908768, 35898917, 60170615, 33557023, 56274520, 33657109, 35595763, 35595855, 18104387, 87169518, 80432113, 22279002, 264564
1317	88055167 (2633, 2634)	Novel Protein sim. GBank g14838757/b1AAD30541.1JAF13491 - (AF134918) semaphorin subclass 4 member G [Mus musculus]	UNCLASSIFIED	264093, 264906, 264900, 264389, 264684
1318	94322893 (2635, 2636)	Novel Protein sim. GBank g14680204/b1AAD27567.1JAF11417 - (AF114171) hypothetical protein [Sorghum bicolor]	UNCLASSIFIED	18108392, 18108348, 265011, 265017, 18108359, 18103682, 56182323, 18108385, 22279000
1319	94238546 (2637, 2638)			264908, 264908, 265006, 265006, 265008, 264592, 265018, 264683, 56181586, 20381171, 29331822, 35595266, 55812038, 265018, 21906768, 265020, 263978, 22279002
1320	86603567 (2639, 2640)	Novel Protein sim. GBank g147401133/b1BAA74870.1] - (AB029554) KIAA0947 protein [Homo sapiens]	UNCLASSIFIED	60432049, 29331828, 264907, 264908, 264909, 264910, 55812038, 264601, 264782, 264784, 264786, 264768, 264709, 264628, 18108374, 264634, 264635, 18108385
1321	86678351 (2641, 2642)	Novel Protein sim. GBank g1486055/b1BAA4337.1] - (AL050276) hypothetical protein [Homo sapiens]	transcript factor	29331828, 264908, 265020, 33657023, 264693, 264404
1322	87755272 (2643, 2644)	Novel Protein sim. GBank g15262591/b1BAA45738.1] - (AL080143) hypothetical protein [Homo sapiens]	RNA_bind	
1323	94845931 (2645, 2646)	Novel Protein sim. GBank g15459516/b1BAA82407.1] - (AB029821) phosphatidylethanolamine N-methyltransferase [Homo sapiens]	synthase	65274572, 56984075, 264259, 29331822, 29331827, 264104, 56182435, 87168474, 18108351, 264283, 21908768, 21908767, 35595917, 265020, 264693, 63274791, 56182323, 18108387
1324	87737614 (2647, 2648)	Novel Protein sim. GBank g15031717/b1BAA05704.1pGBRP - goodpasture antigen-binding protein		22278996, 22278998, 29331828, 264905, 264907, 29331830, 264906, 264510, 265008, 264595, 264759, 21908764, 265018, 264288, 264784, 264786, 264709, 264628, 18108376, 264631, 264632, 264634, 264636, 264638, 264663, 264665, 264666
1325	94447471 (2649, 2650)	Novel Protein sim. GBank g15294501 (U84857) - similar to the DPTK1 unit, family of inhibitors, most similar to tissue factor pathway inhibitor precursor [Caenorhabditis elegans]	protease	35598268, 264905, 264906, 264907, 264908, 264909, 264910, 264593, 33657402, 264758, 85658542, 264760, 264768, 264769, 264691, 35596423
1326	87316289 (2651, 2652)	Novel Protein sim. GBank g11397275 (U81947) - C06C3.8 gene product [Caenorhabditis elegans]	UNCLASSIFIED	264259, 66712552, 264682, 264683, 264635

1327	85522897 (2653, 2654)	Novel Protein sim. GBank Q172852b2p39189ALU2_HUMAN - III ALU SUBFAMILY SB WARNING ENTRY III	Contains protein domain (PF00279) - Plank lipid transfer protein family	UNCLASSIFIED	18108398, 22278996, 22278997, 22278998, 264091, 264259, 2653102, 26531824, 26531825, 26531826, 26531827, 26531828, 264105, 264905, 56182435, 264112, 265008, 265009, 21906754, 265010, 265011, 265017, 265019, 264681, 264446, 264784, 264684, 264288, 264685, 264768, 264686, 21906767, 21906769, 2614823, 2650020, 264690, 264691, 264692, 264693, 263967, 33657109, 33657182, 27486282, 33657349, 18108370, 18108374, 55810764, 36958655, 264634, 26182323, 63373804, 87168518, 60432113, 26182324, 26531825, 26531826, 26531827, 26531828, 26531829, 26531830, 264555, 264488, 22278997, 2633183, 264555, 264567
1328	8753493 (2655, 2656)			UNCLASSIFIED	
1329	8755276 (2657, 2658)	Novel Protein sim. GBank gH487624(gbAAU26969.1)AC00713 - (AC007135) unknown protein sim. GBank gH487624(gbAAU26969.1)AC007135 [Medicago truncatula]		UNCLASSIFIED	22278996, 26531827, 264684, 264692, 33657109
1330	8772737 (2659, 2660)	Novel Protein sim. GBank gH437310 (L23594) - nodulin [Medicago truncatula]		UNCLASSIFIED	264259, 26531825, 264517, 265019, 265021, 264555, 264558, 56528488
1331	8737676 (2661, 2662)	Novel Protein sim. GBank gH4569686(dgJAA76815.1) - (AB023188) KIAA0891 protein [Homo sapiens]		UNCLASSIFIED	264259, 26531826, 26531827, 33666092, 26531828, 60170831, 264448, 264686, 21906765, 55811957, 265020, 33657023, 33657109, 263973, 55811576, 33658422, 33658555, 55182323
1332	8445937 (2663, 2664)	Novel Protein sim. GBank gH55595(gjgJAA82407.1) - (AB029821) phosphatidylethanolamine N-methyltransferase [Homo sapiens]	synthase		60274372, 22278996, 36958655, 22278999, 60432113, 26531827, 26531828, 26531829, 60432289, 26531827, 33658092, 52844045, 55182435, 264510, 21906754, 87168559, 265018, 265019, 264448, 264288, 264369, 264686, 21906765, 21906766, 21906767, 21906768, 265020, 265021, 52844150, 33657023, 55182323, 18108387, 60432113, 22279002
1333	86098476 (2665, 2666)	Novel Protein sim. GBank gH566952(dgJAA83047.1) - (AB023019) KIAA1095 protein [Homo sapiens]		UNCLASSIFIED	60432289, 60712502, 264591, 60433336, 60433438, 55812038, 265010, 264639, 55526486
1334	8759238 (2667, 2668)	Novel Protein sim. GBank gH566259 (AF036680) - Similar to protein-tyrosine phosphatase [Canis familiaris elegans]	phosphatase		22278998, 22278999, 26531827, 264609, 264511, 265007, 265008, 33658095, 60433438, 21906754, 87168559, 265017, 264288, 21906765, 21906767, 21906769, 265018, 265019, 264448, 264288, 264369, 264558, 264638, 264557, 60170394, 264559, 18108385, 264563
1335	8754478 (2669, 2670)	Novel Protein sim. GBank gH424026(dgJAA74921.1) - (AB027053) KIAA0899 protein [Homo sapiens]	B-box zinc finger	UNCLASSIFIED	

1336	8778780 (2671, 2672)	Novel Protein sim. GBank gi464545 p33485 VNUA_PVKA - PROBABLE NUCLEAR ANTIGEN		UNCLASSIFIED	264509, 264505, 264512, 264764, 264693, 264635, 264637
1337	94312042 (2673, 2674)	Novel Protein sim. GBank gi568347 56 BAA3019.1 - (AB028990) KIAA1067 protein [Homo sapiens]		UNCLASSIFIED	56182575, 56994075, 22278938, 22278999, 264092, 264269, 60432289, 29331826, 264906, 264908, 264900, 264112, 265008, 265009, 60433358, 55812038, 33657084, 265011, 265017, 265018, 265019, 264862, 264448, 264863, 264368, 264686, 264695, 21980768, 265008, 60433358, 265011, 265017, 265018, 265019, 1010376, 55811576, 35694855, 56182323, 60432113, 22279002, 265467
1338	80366114 (2675, 2676)			UNCLASSIFIED	26331822, 265010, 264388, 264689, 18108370, 35695555
1339	80249231 (2677, 2678)	Novel Protein sim. GBank gi1178422 (L43194) - rhophilin [Mus musculus]		UNCLASSIFIED	35696032, 264009, 264688, 264550, 264558
1340	88316311 (2679, 2680)				
1341	86101485 (2681, 2682)				264905, 264907, 87166555, 264764
1342	80085017 (2683, 2684)	Novel Protein sim. GBank gi501855 emb CAB44507.1 - (AL036542) d1994E9.5 (h68M-17 (novel 7 transmembrane receptor (rhodopsin family) (oligacy receptor like) protein) [Homo sapiens]	Contains protein domain (PF00001) - Im7 7 transmembrane receptor (rhodopsin family)		264681, 264685, 264686, 264692 264629
1343	90082462 (2685, 2686)	Novel Protein sim. GBank gi457543 ref NP_001384.1 ECM2 - extracellular matrix protein 2	Contains protein domain (PF00560) - struct Leucine Rich Repeat		264910, 264686, 264634
1344	20560560 (2687, 2688)	Novel Protein sim. GBank gi2744101 p 35210 - Ictocarboida cricet. rat (g)			263978
1345	91255546 (2689, 2690)	Ictocarboida cricet. rat (g)		glycoprotein	264908, 60170394
1346	80255717 (2691, 2692)	Novel Protein sim. GBank gi3881052 emb CAA19231 - (AL023643) predicted using GeneFinder similar to serine/threonine kinase; cDNA EST M246a12.3 comes from this gene; cDNA EST M358c10.5 comes from this gene; cDNA EST ENBLM89047 comes from this gene; cDNA EST Y5246a12.5 comes...	Contains protein domain (PF00069) - kinase Eukaryotic protein kinase domain		22278938, 264607, 264681, 264685, 264689, 265020, 264693, 22279000, 22279002, 264566
1347	80417393 (2693, 2694)	Novel Protein sim. GBank gi4504379 ref NP_003658.1 H3C38 - orprian G protein- coupled receptor H3C8	Contains protein domain (PF00560) - glycoprotein Leucine Rich Repeat		264906, 264908, 264909, 265008, 264510, 265011, 265017, 264764, 264766, 264767, 264769, 264631, 264634, 264638, 264697, 264486
1348	87352335 (2695, 2696)	Novel Protein sim. GBank gi3539720 p BAA32100 - (AB010995) peptidylarginine deiminase type IV (Ratus norvegicus)		UNCLASSIFIED	264488, 264489, 264508, 264509, 264510, 264611, 264512, 264591, 264592, 264601, 264604, 264605, 264766, 264532, 264534, 264653, 264536, 264657, 264659, 22278002, 264486

1349	91225448 (2697, 2698)	[Novel Protein sim. GBank g1c14410] [p1j1] 55210 - tricarboxylate carrier - rat (fragment)	UNCLASSIFIED	55646842, 55695384, 22278906, 22278908, 22278909, 264259, 26311122, 26331824, 35660692, 264509, 264509, 264509, 264509, 264509, 264509, 264511, 265008, 264512, 264510, 60170831, 264591, 60433438, 264757, 21906754, 265017, 265018, 264605, 264760, 264762, 264288, 264768, 264889, 21906765, 21906766, 21908757, 21906768, 55811957, 35695917, 265020, 264534, 264691, 264692, 33657023, 264893, 33657349, 18108374, 18108376, 35696423, 60170394, 22279000, 22279002, 264593, 264594
1350	87093136 (2698, 2700)		UNCLASSIFIED	35646842, 264259, 26331825, 264908, 264511, 264884, 264288, 21906769, 265020, 35695917, 264691, 18108374, 35696423, 264555, 264558, 18108385, 22279002, 264466, 264907, 264638
1351	87361327 (2701, 2702)	[Novel Protein sim. GBank g1c48723] g1c48723 (2701, 2702) - (AF084564) BAW protein [Fugu rubripes]	UNCLASSIFIED	
1352	80763686 (2703, 2704)		UNCLASSIFIED	264693, 263981
1353	89345417 (2705, 2706)	[Novel Protein sim. GBank g1c14410] [p1j1] 55210 - tricarboxylate carrier - rat (fragment)	UNCLASSIFIED	35666286, 60424269, 264905, 264509, 264606, 264907, 264908, 264909, 264511, 264512, 264910, 264758, 264596, 55811386, 265011, 264605, 55811150, 264762, 264764, 264766, 52644229, 56181582, 35695917, 265022, 33657023, 264693, 35695763, 60431528, 264629, 263978, 35696423, 35695935, 264630, 264634, 264635, 264636, 264637, 264638, 264639, 18108385, 264583, 264584, 264585
1354	95350845 (2707, 2708)	[Novel Protein sim. GBank g1c488108] g1c488108 (2707, 2708) - hypothetical 43.2 kDa protein [Homo sapiens]	UNCLASSIFIED	22278995, 22278997, 26331826, 264906, 264509, 264512, 264910, 264758, 87188559, 264682, 264388, 264769, 264689, 21906765, 21906766, 52644150, 33657023, 264692, 264693, 18108374, 83373044, 87188518, 22279000
1355	88260186 (2709, 2710)	[Novel Protein sim. GBank g1c14691] g1c14691 (2709, 2710) - (D50928) The KIAA0138 gene product is novel [Homo sapiens]	Contains protein domain (PF00076) - RNA recognition motif. (a.k.a. RRM, RSD, or RNP domain)	22278996, 22278997, 264509, 60714117, 264511, 21906754, 265018, 264769, 264689, 21906765, 21906766, 21906769, 264532, 27486262, 264829, 264638, 264556, 264638, 264639, 264482, 264484

1350	68513981 (2711, 2712)	Novel Protein sm. GBank g1113865 (U40342) - rnm1 [Mus musculus]	sinus	18108337, 22278993, 22278996, 22278998, 264094, 29331826, 264905, 265006, 265007, 265008, 265010, 265017, 265018, 265019, 264764, 18108354, 264689, 21906765, 265022, 18108364, 35696423, 83373044, 18108387
1357	68260268 (2713, 2714)	Novel Protein sm. GBank g16979393jemb CAA903301 - (25020) phosphatidylcholine transfer protein [Bos taurus]	Contains protein domain (PF01652) - START domain	264239, 26331822, 26331825, 264510, 87168559, 266016, 264448, 264286, 21906765, 21906766, 21906768, 265021, 264693, 18108376
1358	38718455 (2715, 2716)	Novel Protein sm. GBank g1656219 (L36831) - transcription regulator [Mus musculus]		264757
1359	87771643 (2717, 2718)		UNCLASSIFIED	264907, 264909, 264510, 264511, 264512, 18108351, 264764, 264534, 33657023, 18108374, 264634, 264635, 264638, 264639, 18108385, 264486, 264567
1360	87738272 (2719, 2720)	Novel Protein sm. GBank g1625822jemb CAA135121 - (Y15417) acylate-CoA ligase [Caprius creticus]	synthase	60432289, 264605
1361	87593327 (2721, 2722)	Novel Protein sm. GBank g16869443jdb BAA83005 11 - (AB028676) KIAA1053 protein [Homo sapiens]	Contains protein domain (PF00536) - SAM domain (Sterile alpha motif)	35696286, 22278997, 22278999, 264259, 26331826, 264506, 264509, 264905, 264907, 264908, 265007, 265009, 33109954, 21906754, 87166474, 265011, 264781, 264683, 264288, 264766, 264769, 264689, 21906768, 265020, 265021, 33657023, 55811576, 35696423, 264634, 60432113, 22279002, 264482, 264486
1362	96287061 (2723, 2724)	Novel Protein sm. GBank g15685411jdb BAA82889 11 - (AB028660) KIAA1037 protein [Homo sapiens]	WD domain, Coata repeat	56162575, 55161666, 60432049, 264259, 26331827, 265012, 265013, 265014, 265015, 265016, 265017, 265018, 265019, 26481208, 16585442, 55811150, 264681, 264288, 264599, 58181582, 60431528, 55810764, 35696423, 60431850, 264558
1363	85758476 (2725, 2726)	Novel Protein sm. GBank g11130434 (L35778) - ADP-ribosylation factor 1-directed GTPase activating protein [Caellus nonnegligus]	Contains protein domain (PF01412) - Putative GTP-ase activating protein 'for Arf'	264488, 26331826, 264607, 264687, 264689, 264693
1364	88179488 (2727, 2728)		UNCLASSIFIED	60432289, 60433356, 60434348, 87168559, 264603, 18108351, 21906766, 35696423, 60432113
1365	83033108 (2729, 2730)	Novel Protein sm. GBank g14688552jdb BAA76803 11 - (AB231715) KIAA0959 protein [Homo sapiens]	oncogene	264766
1366	87003262 (2731, 2732)	Novel Protein sm. GBank g10846443jdb BAA4495 - Hypothetical protein YFR021c - yeast [Saccharomyces cerevisiae]	Contains protein domain (PF00153) - Mitochondrial carrier proteins	265007

1367	6771210 (2733, 2734)	Novel Protein sim. GBank gi464080 enb C4643240.1] - (AL050018), hypothetical protein [Homo sapiens]	Contains protein domain (PF01342) - SAND domain	UNCLASSIFIED	264488, 32546942, 52643355, 22278993, 56994075, 35696286, 22278986, 22278988, 22278992, 264259, 29331824, 66714117, 29331823, 60432289, 35696052, 264905, 264907, 264908, 264909, 264910, 265006, 265007, 264912, 264910, 265009, 60170831, 33657402, 5812038, 21906754, 265011, 21906559, 265017, 265019, 16108351, 264688, 264762, 264683, 264686, 264369, 264690, 264691, 264692, 264693, 264694, 21906766, 21906768, 21906769, 5811957, 265020, 265021, 265022, 264534, 60170815, 264690, 264691, 18108362, 33657023, 33657109, 33657349, 264628, 18108370, 18108374, 18108376, 55811576, 35696423, 35695855, 264635, 264555, 264637, 264556, 52644332, 60170394, 264558, 18108381, 18108385, 56526486, 22279000, 264563, 264567
1368	94320078 (2735, 2736)	Novel Protein sim. GBank gi464561 pp P35389 enb 5_RAT RAS-RELATED PROTEIN RAB-15	Contains protein domain (PF00071) - Ras family	UNCLASSIFIED	264259, 29331822, 29331826, 60432289, 29331827, 35696052, 264508, 264905, 264908, 264908, 264909, 264510, 265007, 264910, 60433436, 264758, 8558542, 8718593, 264600, 264601, 264760, 264764, 264765, 264766, 264767, 264689, 35695917, 265020, 265021, 264632, 264634, 264637, 59844332, 264553, 264632, 63373044, 264563, 264566, 264486, 264597, 265008, 60432229, 60433366, 33657084, 21906764, 21906766, 264555, 264638, 264559, 264557
1370	95316910 (2739, 2740)	Novel Protein sim. GBank gi500182 jpl NP_005823.1 p CKM - potassium large conductance calcium-activated channel, subfamily M, beta member 2		potassium_channel	22278996, 264259, 29331822, 29331824, 29331826, 29331827, 35696052, 264508, 264509, 264905, 264907, 264608, 264909, 264510, 264511, 264512, 264758, 265011, 265019, 264764, 264768, 264769, 21906767, 35695917, 18108362, 35696423, 264632, 264635, 264636, 264555, 264638, 264558, 264639, 18108385, 55274727, 264643, 264643, 264595, 264488
1371	95338512 (2741, 2742)	Novel Protein sim. GBank gi500220 jpl NP_005714.1 p TSPA - tetraspan 5	Contains protein domain (PF00339) - 4 transmembrane segments integral membrane proteins	glycoprotein	29331825, 29331830, 66712902, 265008, 264909, 264758, 33657094, 55585442, 265010, 265018, 265019, 264762, 264448, 35695817, 33657109, 33657182, 33657349, 35695855, 264558, 22279002, 264563

1378	87595071 (2755, 2756)	Novel Protein sim. GBank g9440101 [g9440101-36293] - (A8001772) PEL-S [Cima sangvi]				22278994, 22278998, 264083, 264094, 264259, 26531824, 2631837, 265009, 265018, 265019, 18103351, 264389, 264288, 25148527, 55811957, 264691, 18103396, 33657109, 18103388, 264835, 263381, 18106385
1379	85579344 (2757, 2758)	Novel Protein sim. GBank g9252872 (AF035620) - BRCA1-associated protein 2 [Homo sapiens]			nuc_rept	265020
1380	87627862 (2759, 2760)	Novel Protein sim. GBank g9483737 [g9483737-1] - (AF096834) germ cell specific Y-box binding protein [Homo sapiens]				264510, 264512, 265009, 264288, 264364
1381	88179656 (2761, 2762)	Novel Protein sim. GBank g94731500 [g94731500-1] (AF124538 - (AF123354) LB2A [Drosophila melanogaster])			UNCLASSIFIED	87168559, 265017, 264623, 22279002
1382	94647576 (2763, 2764)		Contains protein domain (PF00543) - Acyltransferase (GNAT) family	UNCLASSIFIED		22278994, 22278997, 22278999, 26531822, 26331834, 26531826, 265007, 60432228, 60433356, 85558542, 265017, 265018, 264685, 264768, 21906766, 35655917, 33657023, 27486261, 27486262, 35655763, 35655855, 87108518, 22279002
1383	87660598 (2765, 2766)			UNCLASSIFIED		18103396, 264692
1384	86915892 (2767, 2768)			UNCLASSIFIED		264488, 264508, 264509, 264505, 264806, 264908, 264909, 264511, 264512, 264910, 264760, 18108351, 264766, 264769, 35655855, 264630, 264630, 264555, 264638, 264483, 264564, 264486
1385	86378788 (2769, 2770)	Novel Protein sim. GBank g92394732 (AF015911) - NAC-1 protein [Rattus norvegicus]			UNCLASSIFIED	35660652, 55811386, 264686, 21906765, 265020, 33657023, 18108355
1386	91013649 (2771, 2772)	Novel Protein sim. GBank g92394910 (AF023982) - contains similarity to the A-type potassium current class of channel proteins [Caenorhabditis elegans]			inf	6042299, 26531828, 264906, 264907, 8162435, 265011, 264681, 60170615, 33657023, 83373044, 264266
1387	87797956 (2773, 2774)	Novel Protein sim. GBank g94160349 [g94160349-1] (AF132162) HSI binding protein 3 [Mus musculus]			UNCLASSIFIED	264931
1388	85101632 (2775, 2776)	Novel Protein sim. GBank g94895164 [g94895164-1] (AF090723) - (AC090723) putative disease resistance protein [Arabidopsis thaliana]			glycoprotein	55274572, 22278999, 264295, 26531826, 26531827, 35699952, 264609, 264907, 264908, 264909, 265006, 265008, 60170831, 33657402, 60433438, 264598, 21906754, 87108559, 264600, 265017, 264683, 18108354, 55644229, 21906765, 21906766, 21906767, 21906768, 21906769, 265021, 264992, 33657023, 33657109, 35655855, 264558, 60170354, 83373044, 22279000

1396	95303233 (2791, 2792)	Novel Protein sim. GBank g 213590 g 154-610 - pH1-ETf1 - human		22276897, 22276999, 264259, 28331825, 60432299, 29331828, 29146498, 29164499, 294307, 264998, 29331830, 264909, 2855006, 265007, 265008, 265009, 60433350, 2855010, 264302, 265017, 265018, 265019, 18108354, 26444228, 18108358, 19106767, 28148627, 265020, 265021, 265022, 265023, 265024, 265025, 265026, 265027, 265028, 265029, 265030, 265031, 265032, 265033, 265034, 265035, 265036, 265037, 265038, 265039, 265040, 265041, 265042, 265043, 265044, 265045, 265046, 265047, 265048, 265049, 265050, 265051, 265052, 265053, 265054, 265055, 265056, 265057, 265058, 265059, 265060, 265061, 265062, 265063, 265064, 265065, 265066, 265067, 265068, 265069, 265070, 265071, 265072, 265073, 265074, 265075, 265076, 265077, 265078, 265079, 265080, 265081, 265082, 265083, 265084, 265085, 265086, 265087, 265088, 265089, 265090, 265091, 265092, 265093, 265094, 265095, 265096, 265097, 265098, 265099, 265100, 265101, 265102, 265103, 265104, 265105, 265106, 265107, 265108, 265109, 265110, 265111, 265112, 265113, 265114, 265115, 265116, 265117, 265118, 265119, 265120, 265121, 265122, 265123, 265124, 265125, 265126, 265127, 265128, 265129, 265130, 265131, 265132, 265133, 265134, 265135, 265136, 265137, 265138, 265139, 265140, 265141, 265142, 265143, 265144, 265145, 265146, 265147, 265148, 265149, 265150, 265151, 265152, 265153, 265154, 265155, 265156, 265157, 265158, 265159, 265160, 265161, 265162, 265163, 265164, 265165, 265166, 265167, 265168, 265169, 265170, 265171, 265172, 265173, 265174, 265175, 265176, 265177, 265178, 265179, 265180, 265181, 265182, 265183, 265184, 265185, 265186, 265187, 265188, 265189, 265190, 265191, 265192, 265193, 265194, 265195, 265196, 265197, 265198, 265199, 265200, 265201, 265202, 265203, 265204, 265205, 265206, 265207, 265208, 265209, 265210, 265211, 265212, 265213, 265214, 265215, 265216, 265217, 265218, 265219, 265220, 265221, 265222, 265223, 265224, 265225, 265226, 265227, 265228, 265229, 265230, 265231, 265232, 265233, 265234, 265235, 265236, 265237, 265238, 265239, 265240, 265241, 265242, 265243, 265244, 265245, 265246, 265247, 265248, 265249, 265250, 265251, 265252, 265253, 265254, 265255, 265256, 265257, 265258, 265259, 265260, 265261, 265262, 265263, 265264, 265265, 265266, 265267, 265268, 265269, 265270, 265271, 265272, 265273, 265274, 265275, 265276, 265277, 265278, 265279, 265280, 265281, 265282, 265283, 265284, 265285, 265286, 265287, 265288, 265289, 265290, 265291, 265292, 265293, 265294, 265295, 265296, 265297, 265298, 265299, 265300, 265301, 265302, 265303, 265304, 265305, 265306, 265307, 265308, 265309, 265310, 265311, 265312, 265313, 265314, 265315, 265316, 265317, 265318, 265319, 265320, 265321, 265322, 265323, 265324, 265325, 265326, 265327, 265328, 265329, 265330, 265331, 265332, 265333, 265334, 265335, 265336, 265337, 265338, 265339, 265340, 265341, 265342, 265343, 265344, 265345, 265346, 265347, 265348, 265349, 265350, 265351, 265352, 265353, 265354, 265355, 265356, 265357, 265358, 265359, 265360, 265361, 265362, 265363, 265364, 265365, 265366, 265367, 265368, 265369, 265370, 265371, 265372, 265373, 265374, 265375, 265376, 265377, 265378, 265379, 265380, 265381, 265382, 265383, 265384, 265385, 265386, 265387, 265388, 265389, 265390, 265391, 265392, 265393, 265394, 265395, 265396, 265397, 265398, 265399, 265400, 265401, 265402, 265403, 265404, 265405, 265406, 265407, 265408, 265409, 265410, 265411, 265412, 265413, 265414, 265415, 265416, 265417, 265418, 265419, 265420, 265421, 265422, 265423, 265424, 265425, 265426, 265427, 265428, 265429, 265430, 265431, 265432, 265433, 265434, 265435, 265436, 265437, 265438, 265439, 265440, 265441, 265442, 265443, 265444, 265445, 265446, 265447, 265448, 265449, 265450, 265451, 265452, 265453, 265454, 265455, 265456, 265457, 265458, 265459, 265460, 265461, 265462, 265463, 265464, 265465, 265466, 265467, 265468, 265469, 265470, 265471, 265472, 265473, 265474, 265475, 265476, 265477, 265478, 265479, 265480, 265481, 265482, 265483, 265484, 265485, 265486, 265487, 265488, 265489, 265
------	-----------------------	--	--	---

1422	88178777 (2843, 2844)	Novel Protein sim. GBank gH495859refNP_000928.1 (pPOLR - polymerase (RNA) II (DNA directed) polypeptide A (220b))		mapolymerase	5694075, 5695986, 8716859, 55411957, 55811576, 264555, 264557, 87168518
1423	86959762 (2845, 2846)			UNCLASSIFIED	264636, 264489, 264692, 264394, 264603, 265018, 264908
1424	93201010 (2847, 2848)	Novel Protein sim. GBank gH437181 (U02289) - GTPase- activating protein [Caenorhabditis elegans]	Contains protein domain (PF00620) - RhoGAP domain	struct	29331822, 29331823, 29331827, 29146484, 264905, 264906, 264908, 264909, 265007, 264910, 265009, 33109354, 265010, 87168559, 285018, 264788, 294887, 21906765, 21906766, 21906767, 21906768, 28148627, 55811957, 29148629, 265021, 264691, 264692, 98526486, 22279002, 284563
1425	21652314 (2849, 2850)	Novel Protein sim. GBank gH100560ipJ514959 - proline- kinase [Homo sapiens]		UNCLASSIFIED	265007, 264558
1426	94322115 (2851, 2852)	Novel Protein sim. GBank gH207844 (L50980) - weak similarity to S. cerevisiae intracellular protein transport protein US11 (SP-P25386) [Caenorhabditis elegans]		UNCLASSIFIED	264448, 60424173, 35695286, 22278907, 22278998, 22278999, 264259, 60419049, 29331822, 29331823, 29331827, 56182435, 264910, 60433356, 60433438, 21906754, 265018, 264288, 21906785, 21906766, 21906767, 21906768, 265020, 265022, 33657109, 18108370, 18108376, 264558, 83373044, 18108385, 56526486, 22279002, 264482
1427	91227510 (2853, 2854)	Novel Protein sim. GBank gH516074gH4AD-5616.1 (AF06194 - (AF06 [943] protease- derived STE20-like kinase PSX [Homo sapiens])	Contains protein domain (PF00069) - Eukaryotic protein kinase domain	kinase	22278894, 56094073, 22278997, 29331828, 29331830, 264628, 265006, 265007, 265008, 265009, 264683, 264288, 18108354, 21906765, 21906768, 29148629, 33657023, 18108374, 35695555, 83373044, 22279002, 284564
1428	94323008 (2855, 2856)	Novel Protein sim. GBank gH18350pp28999VGLX_HSVIEB - GLYCOPROTEIN X PRECURSOR		glycoprotein	28161806, 264259, 264007, 265007, 264909, 264908, 265010, 264946, 4274620, 264628, 63274791, 22279003, 264666
1429	87686669 (2857, 2858)			UNCLASSIFIED	264112, 264558, 265017, 265018, 21906765, 263977, 264555
1430	94735021 (2859, 2860)	Novel Protein sim. GBank gH1181619dJ5A115651 - (D82364) a variant of TSC-22 [Gallus gallus]			264094, 29331824, 264591, 264493, 265018, 264681, 21906765, 21906767, 63274620, 55811576, 264639, 87168518, 22279002
1431	80425081 (2861, 2862)	Novel Protein sim. GBank gH424369jemb.CA46600.1 - (A243460) proteoglycans [Leishmania major]		UNCLASSIFIED	264768, 265020, 264691, 264693, 264638
1432	87630004 (2863, 2864)	Novel Protein sim. GBank gH414797 (L18966) - pyruvate dehydrogenase phosphatase [Bos taurus]		phosphatase	18108394, 29146486, 265007, 60433438, 264763, 29148629, 263969
1433	87605403 (2865, 2866)	Novel Protein sim. GBank gH460316 (AF022147) - uterus ovary specific putative transmembrane protein [Rattus norvegicus]	Contains protein domain (PF00100) - Zona pellucida-like domain	UNCLASSIFIED	264239, 264510, 264591, 264603, 264665
1434	85113730 (2867, 2868)			UNCLASSIFIED	264682, 264691

[illegible]

[illegible]

1464	87620482 (2327, 2929)	Novel Protein sim. GBank gll3974447jmc(CA0272) - (231039) predicted using GeneFinder: cDNA EST EMBL:101209 comes from this gene: cDNA EST y4278a11.3 comes from this gene: cDNA EST y4278a11.5 comes from this gene; cDNA EST y4308a.3 comes from this gene; cDNA EST y4308a.5 com...	UNCLASSIFIED	254569, 22278905, 22278906, 22278907, 22278908, 20331822, 20331824, 20331825, 3568032, 20281100, 264006, 20331830, 264009, 265007, 33657402, 21900754, 265017, 265018, 264682, 264684, 264369, 264388, 264766, 21900765, 21906765, 21906767, 21906769, 35655917, 254691, 33657023, 264692, 35698423, 35695655, 264630, 264631, 264639, 264655
1465	87425192 (2323, 2930)	Novel Protein sim. GBank gll4589598jglj(BAA78821.1) - (A0203194) KIAA0977 protein [Homo sapiens]	glucosylase	254488, 22278964, 56904075, 60432049, 264259, 56182181, 80432269, 29331827, 52644045, 264511, 265007, 265008, 264596, 55872038, 55811386, 264600, 264602, 265017, 265018, 264604, 265019, 18108351, 18108354, 56181562, 21900769, 265021, 33657023, 33657162, 55811576, 264557, 18108359
1468	87606227 (2331, 2932)	Novel Protein sim. GBank gll245532 (U63872) - ORF 73, contains large complex repeat CR 73 [K-ras's sarcoma-associated hepatitis virus]	UNCLASSIFIED	264632, 265017, 264689, 264558
1467	87614320 (2303, 2934)	Novel Protein sim. GBank gll450724 [pe(NP_000137.1) pSSRP - structure specific recognition protein 1]	struct	264683, 264636
1468	89342802 (2933, 2936)	Novel Protein sim. GBank gll450724 [pe(NP_000137.1) pSSRP - structure specific recognition protein 1]		
1469	79236174 (2937, 2939)	Novel Protein sim. GBank gll1906596 (U81788) - kinesin-73 [Drosophila melanogaster]	struct	22278988, 264758, 265018, 265019, 21906769, 265020, 33657109, 22279002
1470	94390482 (2938, 2940)	Novel Protein sim. GBank gll565403jglj(pA0403131.2) F15098 - (A1750992) sly70861 protein [Homo sapiens]	UNCLASSIFIED	18108394, 18108397, 18108398, 32696502, 29160489, 265007, 265008, 265009, 265010, 265011, 18108354, 18108365, 18108388, 18108374, 18108381, 18108382, 18108384, 18108389
1471	87626442 (2941, 2942)	Novel Protein sim. GBank gll397616jgljmc(CA01750) - (Z278542) similar to Mitochondrial carrier proteins. cDNA EST: EMBL:101651 comes from this gene [Caenorhabditis elegans]	Contains protein domain (PF00153) - Mitochondrial carrier proteins	18108394, 18108398, 36182575, 264295, 26331822, 29331824, 29331825, 60432289, 264607, 264909, 265007, 264910, 265008, 265011, 18108354, 18108365, 18108388, 60432438, 264759, 33109954, 265010, 265011, 265018, 264760, 264448, 264764, 264298, 264389, 18108357, 264769, 18108358, 21906767, 21906769, 55811957, 265021, 18108381, 264691, 18108382, 18108385, 18108388, 264628, 18108379, 264637, 264557, 18108381, 56182223, 18108382, 83373044, 18108384, 18108388, 87168518, 60432113, 264404, 22278902, 264482, 264587, 264487
1471	87626442 (2941, 2942)	Novel Protein sim. GBank gll397616jgljmc(CA01750) - (Z278542) similar to Mitochondrial carrier proteins. cDNA EST: EMBL:101651 comes from this gene [Caenorhabditis elegans]	Contains protein domain (PF00153) - Mitochondrial carrier proteins	29331822, 29331824, 29331825, 264828, 264403, 264699, 264693, 18108374, 55811576

[illegible]

1485	34125066 (2371, 2372)	Novel Protein sim. GBank gH589516(djB4376780 1) - (AE023153) KIAA0938 protein [Homo sapiens]	Contains protein domain (PF00056) - Eukaryotic protein kinase domain	kinase	55182575, 22278993, 264696, 264907, 2490594, 168474, 265017, 265019, 1810383, 264696, 265202, 264586
1487	86452711 (2373, 2374)	Novel Protein sim. GBank gH0010275(emh)CAB44331 1 - (AJ132751) xenobiotic/medium-chain fatty acid CoA ligase from XL-III [Bos taurus]		synthase	21906754, 264486
1488	87732026 (2375, 2376)	Novel Protein sim. GBank gH5712131(gBAAQ47379 1AF12046 - (AF120499) DEM1 protein [Homo sapiens]	Contains protein domain (PF01443) - Viral (Superfamily 1) RNA helicase	lyf	264688, 264769, 264689, 264692, 264693, 264509, 264905, 264510, 18108370, 264938, 264623, 264909, 264510, 265006, 264512, 265007, 265008, 265009, 264555, 264556, 264557, 264558, 264762, 264564, 264682
1489	35104277 (2377, 2378)	Novel Protein sim. GBank gH493303p(OZ0708)PFRP_RAT - PROSTAGLANDIN F2-RECEPTOR (PGFR) PRECURSOR (PROSTAGLANDIN F2-ALPHA RECEPTOR ASSOCIATED PROTEIN)	Contains protein domain (PF00047) - immunoglobulin domain	prostaglandin	21906767, 22278993, 265022, 264259, 264693, 26331824, 26331825, 26331826, 26331827, 26331828, 264103, 263972, 66712502, 35696423, 35695855, 265007, 265008, 265009, 83373044, 21906794, 36526486, 265017, 264563, 18108351, 55182575, 264259, 26331827, 26331828, 26331829, 26331830, 26331831, 26331832, 26331833, 26331834, 264905, 66712502, 265007, 265008, 264594, 33657402, 55812038, 87168474, 265018, 18106351, 264369, 264268, 264769, 264689, 21906767, 21906768, 55811957, 60170615, 33657109, 35695955, 264635, 60170394, 55526486, 22279302, 264563
1490	87380127 (2379, 2380)			UNCLASSIFIED	55182575, 264259, 26331827, 26331828, 264598, 66714117, 26331827, 26331828, 264598, 264905, 66712502, 265007, 265008, 264594, 33657402, 55812038, 87168474, 265018, 18106351, 264369, 264268, 264769, 264689, 21906767, 21906768, 55811957, 60170615, 33657109, 35695955, 264635, 60170394, 55526486, 22279302, 264563
1491	83594305 (2381, 2382)	Novel Protein sim. GBank gH255371 (L11275) - selected as a weak suppressor of a mutant of the subunit AC40 of DNA dependent RNA polymerase I and III [Saccharomyces cerevisiae]		UNCLASSIFIED	265007, 264448, 18108372, 264536, 56182323
1492	85805363 (2383, 2384)	Novel Protein sim. GBank gH1656005 (U71205) - rt [Mus musculus]	Contains protein domain (PF00071) - Ras family	oncogene	22279397, 22278990, 26331822, 264907, 66712502

1493	9167215 (2085, 2986)	Novel Protein sim. GBank gl/5689519[gl/BA33041.1] - (A083012) KIA1009 protein [Homo sapiens]	UNCLASSIFIED	264448, 5264365, 63274572, 56182575, 22278994, 55969288, 56994075, 22276599, 60432049, 33331824, 25331828, 35596052, 264508, 264305, 264906, 52644045, 264909, 56182435, 265008, 265008, 265008, 60170831, 33657402, 53812038, 265010, 265011, 265017, 265018, 265019, 53811150, 264448, 264682, 264685, 264686, 52644223, 21906765, 21906766, 21906768, 21906769, 265020, 265021, 60170815, 55944150, 33657023, 18108354, 18108365, 33657108, 33657102, 27468261, 27468262, 27468264, 33657349, 27468265, 33657363, 18108370, 264623, 18108374, 52644332, 56182323, 91748516, 22278002, 264394, 264396, 264397, 265009, 264769, 18108370, 55811576, 2646339, 264665, 264486
1494	87605565 (2087, 2988)	Novel Protein sim. GBank gl/5689519[gl/BA33041.1] - (A083012) KIA1009 protein [Homo sapiens]	Contains protein domain (PF01332) - kinase KRAB box	264448, 5264365, 63274572, 56182575, 22278994, 55969288, 56994075, 22276599, 60432049, 33331824, 25331828, 35596052, 264508, 264305, 264906, 52644045, 264909, 56182435, 265008, 265008, 265008, 60170831, 33657402, 53812038, 265010, 265011, 265017, 265018, 265019, 53811150, 264448, 264682, 264685, 264686, 52644223, 21906765, 21906766, 21906768, 21906769, 265020, 265021, 60170815, 55944150, 33657023, 18108354, 18108365, 33657108, 33657102, 27468261, 27468262, 27468264, 33657349, 27468265, 33657363, 18108370, 264623, 18108374, 52644332, 56182323, 91748516, 22278002, 264394, 264396, 264397, 265009, 264769, 18108370, 55811576, 2646339, 264665, 264486
1495	87605267 (2088, 2990)	Novel Protein sim. GBank gl/5689519[gl/BA33041.1] - (A083012) KIA1009 protein [Homo sapiens]	Contains protein domain (PF01332) - transcription factor KRAB box	22276997, 2642559, 264906, 264607, 265009, 264594, 33657084, 265017, 264760, 284448, 33657109, 264630, 264634, 56526496, 264583, 264585, 264586, 264587, 33669286, 264906, 265019, 264693
1496	87784322 (2091, 2992)	Novel Protein sim. GBank gl/5689519[gl/BA33041.1] - (A083012) KIA1009 protein [Homo sapiens]	UNCLASSIFIED	264910, 264758, 265011, 264764, 264288, 264630, 264634, 264635, 36526496
1497	81695428 (2093, 2994)	Novel Protein sim. GBank gl/5689519[gl/BA33041.1] - (A083012) KIA1009 protein [Homo sapiens]	Contains protein domain (PF00069) - kinase Eukaryotic protein kinase domain	264448, 5264365, 63274572, 56182575, 22278994, 55969288, 56994075, 22276599, 60432049, 33331824, 25331828, 35596052, 264508, 264305, 264906, 52644045, 264909, 56182435, 265008, 265008, 265008, 60170831, 33657402, 53812038, 265010, 265011, 265017, 265018, 265019, 53811150, 264448, 264682, 264685, 264686, 52644223, 21906765, 21906766, 21906768, 21906769, 265020, 265021, 60170815, 55944150, 33657023, 18108354, 18108365, 33657108, 33657102, 27468261, 27468262, 27468264, 33657349, 27468265, 33657363, 18108370, 264623, 18108374, 52644332, 56182323, 91748516, 22278002, 264394, 264396, 264397, 265009, 264769, 18108370, 55811576, 2646339, 264665, 264486
1498	80034338 (2095, 2996)	Novel Protein sim. GBank gl/5689519[gl/BA33041.1] - (A083012) KIA1009 protein [Homo sapiens]	oncogene	264448, 5264365, 63274572, 56182575, 22278994, 55969288, 56994075, 22276599, 60432049, 33331824, 25331828, 35596052, 264508, 264305, 264906, 52644045, 264909, 56182435, 265008, 265008, 265008, 60170831, 33657402, 53812038, 265010, 265011, 265017, 265018, 265019, 53811150, 264448, 264682, 264685, 264686, 52644223, 21906765, 21906766, 21906768, 21906769, 265020, 265021, 60170815, 55944150, 33657023, 18108354, 18108365, 33657108, 33657102, 27468261, 27468262, 27468264, 33657349, 27468265, 33657363, 18108370, 264623, 18108374, 52644332, 56182323, 91748516, 22278002, 264394, 264396, 264397, 265009, 264769, 18108370, 55811576, 2646339, 264665, 264486
1499	86451586 (2097, 2998)	Novel Protein sim. GBank gl/5689519[gl/BA33041.1] - (A083012) KIA1009 protein [Homo sapiens]	glucanase	264448, 5264365, 63274572, 56182575, 22278994, 55969288, 56994075, 22276599, 60432049, 33331824, 25331828, 35596052, 264508, 264305, 264906, 52644045, 264909, 56182435, 265008, 265008, 265008, 60170831, 33657402, 53812038, 265010, 265011, 265017, 265018, 265019, 53811150, 264448, 264682, 264685, 264686, 52644223, 21906765, 21906766, 21906768, 21906769, 265020, 265021, 60170815, 55944150, 33657023, 18108354, 18108365, 33657108, 33657102, 27468261, 27468262, 27468264, 33657349, 27468265, 33657363, 18108370, 264623, 18108374, 52644332, 56182323, 91748516, 22278002, 264394, 264396, 264397, 265009, 264769, 18108370, 55811576, 2646339, 264665, 264486
1500	80495386 (2099, 3000)	Novel Protein sim. GBank gl/5689519[gl/BA33041.1] - (A083012) KIA1009 protein [Homo sapiens]	UNCLASSIFIED	264448, 5264365, 63274572, 56182575, 22278994, 55969288, 56994075, 22276599, 60432049, 33331824, 25331828, 35596052, 264508, 264305, 264906, 52644045, 264909, 56182435, 265008, 265008, 265008, 60170831, 33657402, 53812038, 265010, 265011, 265017, 265018, 265019, 53811150, 264448, 264682, 264685, 264686, 52644223, 21906765, 21906766, 21906768, 21906769, 265020, 265021, 60170815, 55944150, 33657023, 18108354, 18108365, 33657108, 33657102, 27468261, 27468262, 27468264, 33657349, 27468265, 33657363, 18108370, 264623, 18108374, 52644332, 56182323, 91748516, 22278002, 264394, 264396, 264397, 265009, 264769, 18108370, 55811576, 2646339, 264665, 264486
1501	85795267 (3001, 3002)	Novel Protein sim. GBank gl/5689519[gl/BA33041.1] - (A083012) KIA1009 protein [Homo sapiens]	UNCLASSIFIED	264448, 5264365, 63274572, 56182575, 22278994, 55969288, 56994075, 22276599, 60432049, 33331824, 25331828, 35596052, 264508, 264305, 264906, 52644045, 264909, 56182435, 265008, 265008, 265008, 60170831, 33657402, 53812038, 265010, 265011, 265017, 265018, 265019, 53811150, 264448, 264682, 264685, 264686, 52644223, 21906765, 21906766, 21906768, 21906769, 265020, 265021, 60170815, 55944150, 33657023, 18108354, 18108365, 33657108, 33657102, 27468261, 27468262, 27468264, 33657349, 27468265, 33657363, 18108370, 264623, 18108374, 52644332, 56182323, 91748516, 22278002, 264394, 264396, 264397, 265009, 264769, 18108370, 55811576, 2646339, 264665, 264486
1502	80206141 (3003, 3004)	Novel Protein sim. GBank gl/5689519[gl/BA33041.1] - (A083012) KIA1009 protein [Homo sapiens]	transferrase	264448, 5264365, 63274572, 56182575, 22278994, 55969288, 56994075, 22276599, 60432049, 33331824, 25331828, 35596052, 264508, 264305, 264906, 52644045, 264909, 56182435, 265008, 265008, 265008, 60170831, 33657402, 53812038, 265010, 265011, 265017, 265018, 265019, 53811150, 264448, 264682, 264685, 264686, 52644223, 21906765, 21906766, 21906768, 21906769, 265020, 265021, 60170815, 55944150, 33657023, 18108354, 18108365, 33657108, 33657102, 27468261, 27468262, 27468264, 33657349, 27468265, 33657363, 18108370, 264623, 18108374, 52644332, 56182323, 91748516, 22278002, 264394, 264396, 264397, 265009, 264769, 18108370, 55811576, 2646339, 264665, 264486
1503	87012701 (3005, 3006)	Novel Protein sim. GBank gl/5689519[gl/BA33041.1] - (A083012) KIA1009 protein [Homo sapiens]	transferrase	264448, 5264365, 63274572, 56182575, 22278994, 55969288, 56994075, 22276599, 60432049, 33331824, 25331828, 35596052, 264508, 264305, 264906, 52644045, 264909, 56182435, 265008, 265008, 265008, 60170831, 33657402, 53812038, 265010, 265011, 265017, 265018, 265019, 53811150, 264448, 264682, 264685, 264686, 52644223, 21906765, 21906766, 21906768, 21906769, 265020, 265021, 60170815, 55944150, 33657023, 18108354, 18108365, 33657108, 33657102, 27468261, 27468262, 27468264, 33657349, 27468265, 33657363, 18108370, 264623, 18108374, 52644332, 56182323, 91748516, 22278002, 264394, 264396, 264397, 265009, 264769, 18108370, 55811576, 2646339, 264665, 264486

1504	79640061 (2007, 2008)		Novel Protein sim. GBank g1130291[cbjB(A08170)] - (A132545) protein kinase [Homo sapiens]	Contains protein domain (PF00029) - UNCLASSIFIED Ank repeat	264693
1505	66102672 (2009, 2010)		Novel Protein sim. GBank g143373[temp]A141970 [1] - (A132545) protein kinase [Homo sapiens]	Contains protein domain (PF00069) - kinase Eukaryotic protein kinase domain	29331826, 35666032, 264509, 264906, 264907, 264908, 264909, 264511, 264910, 55812038, 264750, 264763, 264764, 264689, 35695937, 265022, 33657109, 18103374, 264631, 264635, 264638, 264566
1506	94143219 (2011, 2012)		Novel Protein sim. GBank g1130291[cbjB(A08170)] - (D26766) alternatively spliced product [Rattus norvegicus]	Contains protein domain (PF00018) - glycoprotein SH3 domain	65274372, 56182575, 56994075, 22278997, 22278998, 22278999, 265091, 264092, 60432945, 264259, 52645080, 29331822, 29331827, 264100, 29331830, 264908, 18108385, 264110, 264511, 264512, 55812038, 264750, 264763, 264764, 264689, 265017, 265018, 264638, 264639, 264689, 264697, 264698, 264699, 264700, 264701, 264702, 21908768, 21908769, 29148629, 52644150, 33657023, 21908770, 26527491, 56182575, 264558, 264559, 18108385, 87166518, 60432113, 22279000, 264565
1507	83736259 (2013, 2014)		Novel Protein sim. GBank g1560851[cbjB(A083040)] - (A029011) KIAA1088 protein [Homo sapiens]	helicase	264639
1508	11619258 (2015, 2016)		Novel Protein sim. GBank g1560851[cbjB(A083040)] - (A029011) KIAA1088 protein [Homo sapiens]	helicase	264593
1509	87319451 (2017, 2018)		Novel Protein sim. GBank g1560851[cbjB(A083040)] - (A029011) KIAA1088 protein [Homo sapiens]	Contains protein domain (PF00069) - kinase Eukaryotic protein kinase domain	264259, 29331822, 29331824, 29331828, 264591, 33109954, 264563
1510	95362643 (2019, 2020)		Novel Protein sim. GBank g1560851[cbjB(A083040)] - (A029011) KIAA1088 protein [Homo sapiens]	UNCLASSIFIED	264596
1511	88318073 (2021, 2022)		Novel Protein sim. GBank g1560851[cbjB(A083040)] - (A029011) KIAA1088 protein [Homo sapiens]	UNCLASSIFIED	264596
1512	95343590 (2023, 2024)		Novel Protein sim. GBank g1560851[cbjB(A083040)] - (A029011) KIAA1088 protein [Homo sapiens]	UNCLASSIFIED	264596
1513	87436228 (2025, 2026)		Novel Protein sim. GBank g1560851[cbjB(A083040)] - (A029011) KIAA1088 protein [Homo sapiens]	UNCLASSIFIED	264596

1514	96345392 (3027, 3028)	Novel Protein sim. GBank gi4593535 p AD23014.1 AC00658 - (AC00658) putative extrinsic suppressor protein (Arabidopsis thaliana)	Contains protein domain (PF01163) - RIO1ZK632 3M,044 family	UNCLASSIFIED	52644607, 52645156, 52646365, 52646842, 55774572, 22276994, 35696285, 56990075, 264259, 52645080, 20331822, 20331925, 35696052, 20331830, 52644044, 56192435, 265006, 60433358, 60433438, 55812038, 21906754, 52646317, 52644296, 87168474, 87168559, 264448, 52644229, 21906765, 21900766, 21906767, 21900766, 35695917, 265020, 52644150, 33657023, 52645129, 33657103, 33657182, 27466261, 27466262, 27468264, 27468265, 35695963, 18103376, 35696423, 35695955, 52644432, 18103385, 18103387, 87168518, 60432113 265020, 264639
1515	79165336 (3029, 3030)	Novel Protein sim. GBank gi3375501 temb CA4377561 - (247812) similar to ubiquitin carboxyl-terminal hydrolase; cDNA EST EMBL D33368 comes from this gene; cDNA EST EMBL D33365 comes from this gene; cDNA EST EMBL D33822 comes from this gene; cDNA EST EMBL D34547 comes from this gene...	ubiquitin	UNCLASSIFIED	265000, 56182323, 22276002
1516	86073539 (3031, 3032)	Novel Protein sim. GBank gi4593535 p AD23014.1 AC00658 - (AC00658) putative extrinsic suppressor protein (Arabidopsis thaliana)	UNCLASSIFIED	UNCLASSIFIED	265000, 56182323, 22276002
1517	87793323 (3033, 3034)	Novel Protein sim. GBank gi3415134 (AF082024) - Phy1	UNCLASSIFIED	UNCLASSIFIED	264091, 18103370, 264404
1518	87350907 (3035, 3036)	Novel Protein sim. GBank gi28538 p P919A ALLB - HUMAN - III ALU SUBFAMILY SK WARNING ENTRY III	hm7	UNCLASSIFIED	66714117, 2645039, 2645089, 264905, 264510, 264910, 264591, 264595, 264288, 264766, 264769, 18103374, 264635, 264638, 264486, 264563, 264489, 60433049, 265009, 33657402, 264596, 21906754, 265019, 264303, 21906765, 21906766, 21906768, 264639, 60433358, 33657184, 27466261, 18103387, 264537, 264639, 87168518, 22276002
1519	94326689 (3037, 3038)	Novel Protein sim. GBank gi5262681 temb CA545771.1 - (AL060198) hypothetical protein (Homo sapiens)	UNCLASSIFIED	UNCLASSIFIED	18103392, 60433049, 264295, 29331824, 265007, 60433356, 265010, 21906768, 264638
1520	87592855 (3039, 3040)	Novel Protein sim. GBank gi2602161 p BAX23712 - (AB077000) HH452 cDNA clone for KIAA0440 has a 438- bp insertion at position 1711 of the sequence of KIAA0440, (Homo sapiens)	Contains protein domain (PF00483) - Nucleoside transferase	UNCLASSIFIED	18103394, 264295, 60714117, 265011, 264603, 265010, 18103394, 33696423, 264557, 264558, 18103388, 29331824, 265018, 265020, 265021 65274572, 21906768, 264693
1521	86970698 (3041, 3042)	Novel Protein sim. GBank gi505235 p AC135616.1 AF135421 GDP- mannose pyrophosphorylase B (Homo sapiens)	UNCLASSIFIED	UNCLASSIFIED	265000, 56182323, 22276002
1522	78606387 (3043, 3044)	Novel Protein sim. GBank gi3776687 (AC003338) - Strong similarity to P21873, gi2800254.1 from A. thaliana BAC 9B/C022560. EST db H65119 comes from this gene, (Arabidopsis thaliana)	UNCLASSIFIED	UNCLASSIFIED	265000, 56182323, 22276002
1523	91003151 (3045, 3046)	Novel Protein sim. GBank gi3776687 (AC003338) - Strong similarity to P21873, gi2800254.1 from A. thaliana BAC 9B/C022560. EST db H65119 comes from this gene, (Arabidopsis thaliana)	UNCLASSIFIED	UNCLASSIFIED	265000, 56182323, 22276002
1524	80203723 (3047, 3048)	Novel Protein sim. GBank gi475904 re NP_004283.1 pR1 - ras inhibitor	UNCLASSIFIED	UNCLASSIFIED	264112, 21906754, 263974, 264683, 264687, 264689, 264690, 264692, 264693
1525	87799887 (3049, 3050)	Novel Protein sim. GBank gi475904 re NP_004283.1 pR1 - ras inhibitor	UNCLASSIFIED	UNCLASSIFIED	264112, 21906754, 263974, 264683, 264687, 264689, 264690, 264692, 264693

1528	95105344 (0051, 3062)	Novel Protein sim. GBank g172850isp0609AMNH, YEAST - GLUCOAMYLASE S152 PRECURSOR (GLUCANT, 4-ALPHA- GLUCOSIDASE) (1:4-ALPHA-D-GLUCAN GLUCOSIDASE) G172850 (YEAST)		glycoprotein	35692266, 56182181, 60431735, 264495, 55012038, 264605, 264683, 2106765, 55011957, 265020, 65274791, 264555, 264556, 264557, 264558, 264559, 83373044
1527	86262512 (0053, 3064)	Novel Protein sim. GBank g1792496 (AF041107) - lup 2 [Rattus norvegicus]			56182375, 264259, 80432040, 26331622, 60432289, 264906, 66717502, 26433438, 21105858, 265017, 264288, 2106766, 264633, 265917, 55811576, 56182323, 18108381
1528	94130918 (0055, 3056)			UNCLASSIFIED	22278953, 22278957, 264259, 66712002, 264598, 265017, 265019, 264682, 264448, 264633, 264764, 264685, 264686, 2106765, 2190706, 21906767, 2106768, 2106769, 265022, 264693, 83373044, 18106385
1529	94120793 (0057, 3058)	Novel Protein sim. GBank g1405663jg0A02053] - (AF131826) Unknown (Homo sapiens)		UNCLASSIFIED	264488, 263994, 56182575, 22278953, 35692266, 22278957, 264259, 26331822, 60432289, 26331827, 35696052, 264509, 264906, 264907, 264908, 264909, 52644045, 56182435, 264511, 265009, 264910, 60433396, 60433438, 265017, 265018, 264760, 264448, 264764, 264389, 264288, 264766, 18106337, 264768, 52644229, 265019, 21906766, 21906767, 21906768, 265017, 265018, 265019, 265020, 265021, 264629, 35695845, 80432111, 22278902, 264563, 264564, 264486, 264567
1530	55012765 (0059, 3060)	Novel Protein sim. GBank g12828710 (AF043642) - malin cyclophilin [Rattus norvegicus]			264488, 264489, 35695266, 26331925, 35696052, 264508, 264905, 264906, 264907, 264908, 264510, 264511, 264512, 264910, 264592, 264595, 18106331, 264764, 264683, 264684, 264766, 264768, 18106337, 264769, 35695917, 264628, 264629, 18106374, 35695955, 264630, 264631, 264634, 264555, 264636, 264637, 264404, 264563, 264588, 264466
1531	95413351 (0061, 3062)	Novel Protein sim. GBank g11935674 (J50878) - catboxyl terminal LIM domain protein (Homo sapiens)		Kinase P42 domain (Also known as DHR or GLGF).	56182375, 35696052, 264507, 264259, 26331822, 26331825, 26331826, 26331827, 35696052, 264509, 56182435, 264510, 264511, 265007, 60433396, 55811986, 264509, 264510, 264511, 264512, 264513, 55811987, 35696057, 35672723, 35695763, 55810764, 35696423, 55811578, 263981, 5070394, 56182323, 63373044, 60432113, 264565

1532	65716224 (3063, 3064)	Novel Protein sim. GBank gl 38147476 tm0j CAJ91256 - (Z6494) cDNA EST EMBL:D68271 comes from this gene; cDNA EST EMBL:D64490 comes from this gene; cDNA EST EMBL:D64490 comes from this gene; cDNA EST EMBL:D67438 comes from this gene; cDNA EST EMBL:D68087 comes from this gene; cDN			UNCLASSIFIED	264699
1533	94239830 (3065, 3066)	Novel Protein sim. GBank gl 1460324 tm0j CAJ91543 - (Z78141) unknown [Mus musculus]			silud	29331824, 2616409, 264907, 264112, 265006, 265011, 265017, 265018, 264762, 18108351, 263987, 20231149, 18108374, 263981, 264560
1534	55343941 (3067, 3068)	Novel Protein sim. GBank gl 81280 pt S22897 - extensin -			UNCLASSIFIED	264905, 264907, 264766, 264637
1535	50938732 (3069, 3070)	Velox carter (fragment)				65274572, 22278697, 264259, 60432049, 29331822, 60432289, 29331827, 2918459, 265006, 265008, 60170831, 60433438, 33109954, 67168559, 265018, 18108357, 21907766, 29148629, 265021, 265022, 18108377, 56182323, 60432113, 22279000, 22279002
1536	87602856 (3071, 3072)	Novel Protein sim. GBank gl 106623 pt B32651 - finger protein 2, plectin - human	Contains protein domain (PF00066) Zinc finger, C2H2 type		transcript factor	264696, 18108357, 18108394, 21907767, 21907768, 29148629, 35646286, 265020, 265021, 35644150, 264693, 66714117, 29331822, 60432289, 29331827, 2918459, 265006, 265008, 60170831, 60433438, 20231149, 264909, 18108377, 5696423, 35955655, 265009, 264634, 264636, 264638, 18108385, 56528488, 285017, 285018, 264593, 264762, 18108351, 264448, 264369, 264766
1537	96354556 (3073, 3074)	Novel Protein sim. GBank gl 3876332 tm0j CAJ92056 - (Z79754) cDNA EST EMBL:T01054 comes from this gene; cDNA EST EMBL:D73500 comes from this gene; cDNA EST yk42612.5 comes from this gene; cDNA EST yk42610.5 comes from this gene; cDNA EST yk47505.5 comes from this gene; cDNA ES...				65274572, 56182575, 60432048, 264259, 29331826, 265006, 265007, 60433356, 60433438, 264601, 18108351, 264448, 264369, 264488, 33657023, 65274620, 33657106, 60432113
1538	65724628 (3075, 3076)	Novel Protein sim. GBank gl 403340 tm0j 787 - [Gallus domesticus skeletal muscle mRNA, partial cds], gene product [Gallus gallus]	Contains protein domain (PF00069) Eukaryotic protein kinase domain		kinase	18108394, 18108397, 264909, 265008, 265009, 265010, 18108351, 264638, 18108382, 18108385, 18108388

[illegible]

1548	B4233065 (2005, 2006)	Novel Protein sim. GBank [g 3034892 gdb BAA25510] - (A8011155) KIA0054 sim. [Homo sapiens]	UNCLASSIFIED	pH	29331824, 60431528, 264635, 56182323
1549	B5330048 (2007, 2008)	Novel Protein sim. GBank [g 558515 gdb BAA30433] - (A8028014) KIAA1091 Protein [Homo sapiens]	UNCLASSIFIED		60424179, 22278995, 35589036, 22278988, 22278999, 264092, 264094, 29331822, 56182181, 29331824, 35596052, 264905, 264906, 264907, 264909, 265006, 264511, 265008, 60437335, 60433356, 21906754, 5581398, 87185559, 265017, 265018, 265019, 55811150, 264682, 264288, 264369, 56181562, 264789, 21906765, 21906768, 21906769, 55811857, 565020, 264691, 565018, 60431528, 35596043, 35685855, 56525466, 60435123, 22279002, 264563, 264566, 264568
1550	95201907 (2009, 2010)	Novel Protein sim. GBank [g 544463 gdb P35350 GUSEL - POSSIBLE GUSTATORY RECEPTOR TYPE 6 (P6RI PROTEIN)]	UNCLASSIFIED	7 transmembrane protein domain (PF00001) - 7 transmembrane receptor (rhodopsin family)	65274574, 60432389, 265008, 264910, 265011, 265017, 264768, 56182323
1551	88077111 (2101, 2102)	Novel Protein sim. GBank [g 4735566 gdb NP_004788.1 p I565 - Hepatocellular carcinoma sulfotransferase	UNCLASSIFIED		22278989, 26331822, 264508, 264509, 264806, 264907, 264909, 265007, 264512, 264910, 21906754, 255018, 265019, 264681, 264764, 264766, 264688, 264769, 21906769, 264692, 35695763, 264635, 264555, 264556, 264557, 264838, 264558, 264563
1552	87617114 (2103, 2104)	Novel Protein sim. GBank [g 458970 gdb BAA76007] - (A8023100) KIA0469 protein [Homo sapiens]	UNCLASSIFIED		264259, 26331828, 56712502, 264764, 264288, 264686, 33657109, 264558
1553	94725512 (2105, 2106)	Novel Protein sim. GBank [g 458970 gdb BAA76007] - (A8023100) KIA0469 protein [Homo sapiens]	UNCLASSIFIED	-dehydrogenase Gamma-thionins family	56182575, 35596286, 29316499, 264509, 264907, 264908, 265009, 56182435, 265006, 265008, 265009, 264910, 264757, 264758, 265017, 55811150, 18108351, 264764, 56181562, 35695917, 264693, 33667109, 18108374, 35696423, 63274791, 35695955, 264635, 264555, 56182323, 18108382, 83373044, 22277990
1554	94323069 (2107, 2108)	Novel Protein sim. GBank [g 3034892 gdb BAA25510] - (A8011156) KIA0054 protein [Homo sapiens]	UNCLASSIFIED	transferrase Gonadotropin-releasing hormones	35595928, 22277990, 264259, 26331822, 29331824, 29331825, 29331828, 265007, 265009, 60432229, 33657402, 55812038, 265011, 265019, 264681, 264395, 264686, 264767, 264768, 21906765, 21906769, 35695917, 264692, 18108351, 60431528, 55811578, 264683, 60129044, 56181562, 83373044, 18108385, 22293900, 22729202

1555	87332970 (3109, 3110)	Novel Protein sim. GBank g12257456[BLA241332]- (A004354) p015 (Schistosoma mansoni proteid)	Contains protein domain (PF00400) - WD domain, C-termin repeat	UNCLASSIFIED	264259, 26331826, 35695052, 264508 264305, 264096, 264007, 264001, 2644045, 264909, 264910, 60432203, 60433596, 26417038, 264758, 264759, 33657394, 265011, 87168559, 264601, 265016, 265019, 264763, 264764, 264288, 264765, 264768, 21906765, 35695917, 265022, 264891, 33657023, 35695943, 35695555, 264635, 264555, 264636, 264638, 264639, 18108335, 56526486
1556	91225268 (3111, 3112)			UNCLASSIFIED	83373044, 264758, 265022, 264600, 35966052, 264630, 35696643, 265018, 264632, 264682, 29331822, 265020, 265011
1557	87640609 (3113, 3114)	Novel Protein sim. GBank g1329611 (AF078783) - contains similarity to C3HC4-type zinc fingers (Pfam; zf- C3HC4 Hmm. score: 34.08); most similar to discophila novel protein (SW: 006003) [Caenorhabditis elegans] Golang Protein sim. GBank g13300105[p]A042871 (1AF15510 - (AF155105) putative zinc finger protein NY-REN-34 antigen (Homo sapiens)	Contains protein domain (PF00097) - Zinc finger, C3HC4 type (RING finger)	UNCLASSIFIED	264909, 264910, 264768, 264769, 264781, 264782, 264786, 264789, 264690, 263978, 264634, 264635, 264639, 264564, 264466 22278594, 22278596, 22278697, 22278698, 22278699, 60432046, 264459, 29331824, 29331825, 29331826, 29331827, 264908, 264909, 60433356, 21900754, 265017, 265018, 264448, 21906767, 265021, 265022, 33857023, 33657109, 18108370, 55811576, 83373044, 87168518, 22279600, 22279002 264259, 29331822, 60432389, 35696052, 264107, 264110, 21900754, 33108954, 87168559, 264780, 264763, 21900764, 21906765, 21906769, 265021, 264690, 35959585
1559	86224665 (3117, 3118)	Novel Protein sim. GBank g11129009[gp]027100/A2GL - HUMAN - L-LEUCINE-RICH ALPHA-2 GLYCOPROTEIN (LRG)	Contains protein domain (PF00500) - glycoprotein Leucine Rich Repeat	UNCLASSIFIED	264908, 264603, 264638
1560	84590675 (3119, 3120)	Novel Protein sim. GBank g13880146[em]CA562704 - (Z68319) Similarity to Human hRNP F protein (P/R Acc. No. S43464); cDNA EST EMBL D34218 comes from this gene; cDNA EST EMBL D37248 comes from this gene; cDNA EST EMBL D71817 comes from this gene; cDNA EST EMBL D74531 comes fro...		UNCLASSIFIED	
1561	86069159 (3121, 3122)			UNCLASSIFIED	264510, 264594
1562	83359652 (3123, 3124)			UNCLASSIFIED	264910, 264754, 264766
1563	85596894 (3125, 3126)			UNCLASSIFIED	18103394, 35695856, 264656, 29331822, 80432289, 35696052, 29331928, 264608, 86715202, 264698, 69182433, 265007, 264910, 60170831, 21906754, 265011, 265017, 265018, 265019, 18108351, 264448, 264288, 264686, 21906765, 21900768, 265021, 60170815, 264692, 35696843, 35695855, 264557, 56182323, 60421213, 22279002, 264482
1564	87766371 (3127, 3128)	Novel Protein sim. GBank g11188207[gp]495304/CDV RAT - ACYL-CoA DEHYDROGENASE VERY-LONG-CHAIN SPECIFIC PRECURSOR (VLCD)	Contains protein domain (PF00441) - Acyl-CoA dehydrogenase	UNCLASSIFIED	

1574	54216142 (3147, 3148)	Novel Protein sim. GBank g175334p1pNP_004258.1pRADS_delta-6 fatty acid desaturase	Heme-binding domain in cytochrome b5 and oxidoreductases	Cytochrome	18108394, 264887, 18108397, 18108398, 2331825, 29331827, 29146490, 29146499, 264107, 264807, 264809, 52444045, 264511, 265038, 264910, 265009, 264531, 21906754, 265011, 265011, 265011, 265011, 265011, 2647954, 18108354, 264389, 264288, 264786, 264698, 264698, 21906765, 21906766, 21906767, 21906768, 21906769, 29148829, 264690, 264691, 264693, 20281069, 18108370, 18108374, 18108379, 35695855, 264634, 18108384, 18108385, 22279002, 264583, 264586
1575	95340019 (3149, 3150)	Novel Protein sim. GBank g13861810jemb1CA948561 - (Z70783) similar to EF-hand calcium binding protein; cDNA EST EMBL C06700 comes from this gene [Caenorhabditis elegans]	Contains protein domain (PF00039) - EF hand	phosphatase	29331825, 264529, 29331822, 29331824, 265006, 265008, 265009, 265010, 87168559, 55811150, 264448, 18108354, 264389, 264288, 18108357, 55811957, 265020, 265021, 60170615, 264691, 33857023, 33857108, 60431526, 63274791, 35695855, 18108395, 60431173, 22279002, 264586
1576	95314019 (3151, 3152)	Novel Protein sim. GBank g1273105 (AF039711) - contains similarity to Phycocyanin proteins (gamma-cyridide 3-phosphate dihydrogenase (GBX72381) [Caenorhabditis elegans])	UNCLASSIFIED	UNCLASSIFIED	264506, 264905, 264509, 264907, 264909, 264510, 264511, 265006, 264910, 21906754, 265010, 265011, 87168559, 264791, 264762, 264288, 264795, 264799, 264691, 264693, 35695855, 264632, 264634, 264635, 264638, 83373044, 264486
1577	87613800 (3153, 3154)	Novel Protein sim. GBank g12489130pP70315WASP_MOUSE - WISKOTT - ALDRICH SYNDROME PROTEIN HOMOLOG (WASP)	UNCLASSIFIED	UNCLASSIFIED	264511, 265011, 264681, 264389, 264586, 264889, 264629, 264555, 264559, 264559
1578	87123138 (3155, 3156)	Novel Protein sim. GBank g12489130pP70315WASP_MOUSE - WISKOTT - ALDRICH SYNDROME PROTEIN HOMOLOG (WASP)	UNCLASSIFIED	UNCLASSIFIED	264259, 29331826, 265017, 264699, 264693, 60432113
1579	88085141 (3157, 3158)	Novel Protein sim. GBank g1278255p1BAA251901 - (A6007407) myeloid zinc finger protein-2 [mus musculus]	Contains protein domain (PF00396) - Zinc finger, C2H2 type	UNCLASSIFIED	35695828, 264908, 264909, 60433438, 55811396, 264389, 264685, 33857023, 264585, 264585, 264587, 97168518, 264586, 264586, 264586, 264586, 55123038, 87168471, 265018, 264444, 265022, 264638, 55826486, 22279000
1580	87253702 (3159, 3160)	Novel Protein sim. GBank g13248230p1A4D168661 - (AF108674) late gestation lung protein 1 [Rattus norvegicus]	Contains protein domain (PF00188) - SCP-like extracellular protein	glycoprotein	22278961, 29331822, 26531824, 29331826, 55182433, 264595, 55812038, 87168559, 265017, 264288, 21906784, 55811957, 35695917, 264692, 55811578, 264637, 55182323, 264559, 83373044, 26432113
1581	95087431 (3161, 3162)	Novel Protein sim. GBank g12038838 (AF003386) - F59E12.4 gene product [Caenorhabditis elegans]	UNCLASSIFIED	UNCLASSIFIED	

1582	95350052 (3163, 3164)	Novel Protein sim. GBank g1542037emj(CAB46079.1) - (A424459) proteophosphoglycan [Leishmania major]		phosphatase	264253, 60432288, 29331827, 2645058, 264905, 264906, 264907, 264909, 264910, 264762, 264288, 264765, 264915, 264632, 264935, 264939, 264946, 22278980, 264955, 264962, 264662, 21900768, 15595845, 264463
1583	87622715 (3165, 3166)	Novel Protein sim. GBank g15576956temj(CA351131.1) - (A050206) 5,4,7,9,17,2 (novel protein) [Homo sapiens]		UNCLASSIFIED	89042179, 52648652, 65747677, 56192575, 22278980, 36968296, 22278996, 22278998, 22278999, 264256, 29331822, 46182181, 60424269, 60432289, 29331827, 29331828, 35696052, 29146488, 66712502, 29331830, 52644045, 56182435, 264510, 264512, 265008, 60433356, 33657402, 60433438, 55612038, 21906754, 55611386, 52644296, 87168474, 87168559, 265018, 265019, 264448, 264369, 264298, 18100358, 21906765, 21906767, 21906768, 21906769, 35695937, 265020, 265021, 265022, 52644150, 33657023, 33657109, 18108374, 55610764, 55611576, 55616523, 66742371, 21906768, 21906769, 33373044, 8100387, 87168474, 60432113, 22278998, 22278999, 22278998, 264256, 264905, 265006, 29331824, 29331825, 264905, 33109954, 265007, 265008, 60433356, 33109954, 87168474, 265011, 265017, 2646004, 264369, 264288, 264685, 264769, 18108359, 21906765, 18108364, 18108370, 264029, 263972, 18108383, 18108388, 264482, 264554
1585	87626117 (3169, 3170)	Novel Protein sim. GBank g142401320u(BA174846.1) - (A80026500) KIAA0823 protein [Homo sapiens]	Contains protein domain (PF00023) - Anki repeat	phosphatase	35696246, 22278998, 264256, 265006, 265007, 265008, 60433356, 33109954, 87168474, 265011, 265017, 2646004, 264369, 264288, 264685, 264769, 18108359, 21906765, 18108364, 18108370, 264029, 263972, 18108383, 18108388, 264482, 264554
1586	86067081 (3171, 3172)	Novel Protein sim. GBank g13786454 (AF098393) - No definition line found [Caenorhabditis elegans]		UNCLASSIFIED	265017, 265018, 264689, 33657023, 263978, 264636, 264653
1587	87617126 (3173, 3174)	Novel Protein sim. GBank g13253159 (AF008385) - translation initiation factor eIF2C [Oryzias latipes]		UNCLASSIFIED	264907, 264908, 264411, 264910, 264951, 264994, 264825, 264631, 264583, 264463, 264567
1588	87602536 (3175, 3176)	Novel Protein sim. GBank g10775737p(S25890 - probable ribosomal protein L34, mitochondrial - yeast [Saccharomyces cerevisiae])	Contains protein domain (PF00468) - Ribosomal protein L34	UNCLASSIFIED	294235, 29331828, 264695, 265008, 264758, 264764, 264765, 264782, 21906765, 21906769, 60170615, 52644150, 33657109, 35695845, 56182233, 18108385, 65747677, 264480, 29331822, 86714117, 29331827, 29331828, 56182435, 265008, 60170831, 264595, 264758, 264586, 265011, 264686, 21906765, 21906768, 55611957, 27486265, 264636, 18108385, 56526486, 60432113
1589	90956563 (3177, 3178)	Novel Protein sim. GBank g12137756p(IL48746 - semaphorin C - mouse (fragment))		UNCLASSIFIED	

1590	95318825 (3178, 3180)				UNCLASSIFIED	264489, 22278996, 264250, 26331824, 29331825, 26331826, 26331827, 2650008, 69433366, 21906754, 2650118, 2650119, 264448, 264765, 264286, 52844220, 2650119, 264448, 264765, 264286, 52844220, 21906754, 21906757, 21908788, 21906789, 265001, 264692, 27480255, 35695763, 56526496, 60432113, 22279000, 22279002, 264584
1591	96877160 (3181, 3182)	Novel Protein sim. GBank			MHC	264259, 284805, 29331830, 284855, 265017, 264448, 264288, 264690, 264620, 97168518
1592	97882533 (3183, 3184)	g14557749e[NP_000237.1pMHC2 - MHC class II transactivator				
1593	94991661 (3185, 3186)				UNCLASSIFIED	55274572, 50432049, 264509, 50433358, 21906754, 21908767, 21906788, 18108370, 35696423, 22279000, 264595, 264567
1594	97773732 (3187, 3188)	Novel Protein sim. GBank g18877072[emb[CMA87060] - (246937) similarity with ribosomal protein L21 [Caenorhabditis elegans]			UNCLASSIFIED	264483, 29331827, 264903, 264906, 264807, 264908, 264909, 264910, 264592, 264993, 264757, 264502, 264604, 264760, 264881, 264288, 264765, 264768, 29148029, 35695917, 264592, 264628, 264620, 264630, 264632, 264634, 264635, 264636, 264639, 264593, 264584, 264586
1595	79519425 (3189, 3190)	Novel Protein sim. GBank g13152703 [AF053599] - (14899) NCT-1 [Homo sapiens]			UNCLASSIFIED	29331826, 264698, 55811957
1596	76343828 (3191, 3192)	Novel Protein sim. GBank g12527114p[AA001424.1(AF094448 - (AF094480) Cholesterol 24-hydroxylase [Homo sapiens]			UNCLASSIFIED	29314698, 264758, 2653867
1597	96971857 (3193, 3194)				cyto450	264032, 29331824, 264506, 264982, 264366, 264686, 264630, 264593
1598	37602939 (3195, 3196)					252959, 264634
1599	87643828 (3197, 3198)	Novel Protein sim. GBank g14987679[pe[NP_000324.1]pSCA7 - spinocerebellar ataxia 7 (homoautosomal atrophy with retinal degeneration)			UNCLASSIFIED	52845000, 29331823, 26331826, 264511, 265009, 265011, 264605, 264448, 264764, 265020, 264682, 264693, 18109370, 204635, 18109385
1600	80556002 (3199, 3200)					29331826, 264693, 264691, 264593
1601	15023246 (3201, 3202)	Novel Protein sim. GBank g13057049p[AA017861.1(AF12853 - (AF128535) Cytosolic phosphogluconate (NAD ⁺)-dependent) chicken			UNCLASSIFIED	264635
1602	86929887 (3203, 3204)				stud	29146499, 264112, 264762, 18109351, 29146627, 263974
1603	80502072 (3205, 3206)				collagen	264440, 29331824, 264607, 264608, 264511, 265009, 264592, 265010, 265011, 264762, 264764, 264598, 264288, 264687, 264768, 264763, 264628, 264634, 264635, 264593, 264638, 264597, 264596, 264599, 18109385
1604	90221813 (3207, 3208)	Novel Protein sim. GBank g14758833[pp[AA029633.1(AF11682 - (AF116827) unknown [Homo sapiens]			ATPase-associated	263977

1621	87075708 (3241, 3242)	Novel Protein sim. GBank g 3822553 (AF080788) - nuclear calmodulin-binding protein [Gallus gallus]	UNCLASSIFIED	264910 18100393, 65274572, 18108398, 22278996, 22278997, 22278999, 29146498, 29146499, 264905, 264908, 264909, 264928, 52644045, 264592, 60433556, 21906754, 264602, 265017, 264369, 21906768, 55811957, 265021, 60170515, 264435, 264557, 60170394, 83373044, 18108385, 22279000, 22279002, 264456
1622	87779106 (3245, 3246)	Novel Protein sim. GBank g 7310898 (P03090V42, SORHO - UN-INDUCED PROTEIN LM22)	ribosomalprot	18108398, 264459, 264909, 56181435, 87169474, 264448, 21906768, 35695917, 264691, 87169518, 264563
1624	87338178 (3247, 3248)	Novel Protein sim. GBank g 3875566 (p3) (A0954781 - (231104) cDNA EST EMBL D03665 comes from this gene; cDNA EST EMBL D03654 comes from this gene; cDNA EST y-2408.3 comes from this gene; cDNA EST y-387c8.3 comes from this gene; cDNA ES...	UNCLASSIFIED	86714117, 29331825, 264909, 265008, 264758
1625	95354748 (3249, 3250)	Novel Protein sim. GBank g 4586822 (p5) (BAA76833.1) - (A8023206) KIAA0989 protein [Homo sapiens]	kinase	364489, 22278994, 22278995, 22278996, 35695916, 22278997, 22278998, 22278999, 264092, 264259, 29331824, 29331825, 29331827, 29331828, 264102, 264106, 264506, 33657084, 265017, 265018, 18108351, 264683, 264389, 264288, 21906765, 21906766, 21906767, 21906769, 35695917, 265021, 264691, 65274620, 18108368, 263972, 18108376, 35696423, 264631, 264434, 22279000, 22279002
1626	94734389 (3251, 3252)	Novel Protein sim. GBank g 6790708 (BACD4884.1) (AF16090 - (AF160904) cDNA HL05936 [Drosophila melanogaster])		52644507, 52645156, 52646365, 52646842, 22278994, 59964075, 22278997, 22278998, 22278999, 60432049, 264259, 29331822, 29331824, 60714117, 29331825, 60432289, 29331827, 29331828, 35696052, 604906, 604907, 604908, 604909, 604910, 604911, 604912, 604913, 604914, 604915, 604916, 604917, 604918, 604919, 604920, 604921, 52646317, 52644296, 265011, 87168569, 264604, 265018, 264448, 264389, 264288, 264766, 52644239, 264689, 21906765, 21906768, 35695917, 265021, 265022, 52644150, 33657023, 65274620, 27486261, 27486262, 27486265, 35695763, 263972, 52644332, 60170394, 87168518, 60432113, 264567
1627	83368173 (3253, 3254)	Novel Protein sim. GBank g 3646037 (AC004657) - hypothetical protein [Arabidopsis thaliana]	UNCLASSIFIED	264288
1628	85708459 (3255, 3256)		epn	264288, 264686, 264767, 22279002

[illegible]

1640	94143185 (3279, 3280) Novel Protein sim. GBank gi 2842469 jens(CAA16847.1) - (AL021747) hypothetical protein [Schistosoma haematophyllum]	UNCLASSIFIED	56182575, 56994075, 35658586, 60433049, 56474755, 56474755, 56474755, 56474755, 56474755, 264510, 265006, 265007, 265008, 264810, 265009, 33657402, 55812038, 265010, 265011, 265017, 265018, 264288, 52644220, 21906785, 21906788, 21906788, 31695917, 265021, 60170615, 52044150, 33657023, 33657109, 33657348, 18108374, 35696423, 65274791, 35695855, 264632, 264555, 58182323, 22279000
1641	87625160 (3281, 3282)	UNCLASSIFIED	29148489, 265006, 265007, 265009, 265009, 55812038, 265010, 265011, 264555, 264556, 264558, 18108383
1642	94312557 (3283, 3284) Novel Protein sim. GBank gi 1575333 (U90416) - myr 6 myosin heavy chain [Rattus norvegicus]	Contains protein domain (PF01843) - striat DII domain	22278999, 29147620, 26531826, 29331828, 33659970, 55812038, 265010, 265018, 265019, 18108351, 264689, 265020, 265022, 264690, 33657023, 65274620, 35695783, 32695324, 61686161, 61686162, 61686163, 18108388, 87168518, 22279002, 264564, 26531825, 26531827, 26531828, 21906754, 265014, 264288, 264689, 33657349, 18108370, 18108376, 264555, 83373044, 22279002, 264482
1644	85026125 (3287, 3288)	UNCLASSIFIED	264805, 264907, 264908, 264910, 265009, 264757, 264758, 264761, 264762, 264763, 264766, 264768, 264769, 264628, 264629, 264630, 264631, 264632, 264633, 264564, 264555, 264566, 264567
1645	55013558 (3289, 3290) Novel Protein sim. GBank gi 1076020 jens(S49915) - extensin like protein - maize	UNCLASSIFIED	22278994, 59994075, 35696266, 294259, 26331824, 29331825, 26331826, 00422289, 264508, 00433356, 00433438, 87106559, 263018, 264697, 35695917, 264692, 33657023, 35695782, 265010, 264563, 264564, 264565, 264566, 264567
1647	94278428 (3293, 3294) Novel Protein sim. GBank gi 5050237 jens(C844338.1) - (Y17465) alpha-N-acetylglucosaminase alpha 2-B- glucosyltransferase [Cigu ranae]	UNCLASSIFIED	26531823, 264908, 264908, 264909, 21906768, 60170615, 264639, 22279000
1648	87642098 (3295, 3296)	Contains protein domain (PF00098) Zinc finger, C2H2 type	265009, 264686, 55811957, 35695917, 55810764, 264556, 50182323, 264558, 18108385

1649	95347628 (3297, 3298)	Novel Protein sim. GBank gi 854065 emb CA438337 - (X43413) U88 [Human herpesvirus 6]	cadherin	264488, 22278995, 35950286, 22278996, 22278997, 22278999, 60432048, 264259, 29331822, 29331824, 29331825, 29331826, 29331827, 29331828, 264906, 264907, 29331829, 264908, 264909, 264910, 264911, 265007, 265008, 265009, 264991, 264993, 60433438, 264596, 55912038, 21906754, 265011, 264607, 264602, 265017, 265018, 265019, 264682, 264448, 264764, 264683, 264258, 264766, 264685, 264687, 264768, 264688, 264769, 52644229, 264689, 21906765, 21906768, 21906767, 21906768, 55811957, 35655917, 265021, 265022, 52644150, 264682, 33857023, 33657109, 20281148, 18108370, 264628, 18108374, 18108378, 35696423, 35695855, 264632, 264634, 264635, 264638, 18108380, 264639, 264638, 18108382, 18108384, 18108385, 18108387, 264680, 264404, 60432113, 22278900, 22278902, 264482, 264595, 264596, 264597, 264598, 21906767, 18108374, 18108377, 18108385
1650	87418539 (3299, 3300)	Novel Protein sim. GBank gi 564735 emb CA421059 - (X431644) possible zinc finger protein		
1651	91639773 (3301, 3302)	Novel Protein sim. GBank gi 4884278 emb CA43247.1 - (AL050037) hypothetical protein [Homo sapiens]	synthase	264488, 52645156, 18108397, 35696286, 22278998, 22278999, 264249, 29331822, 29331824, 29331825, 29331826, 29331827, 29331828, 264508, 264908, 29331830, 264910, 60432226, 21906754, 265010, 265011, 265017, 265019, 264448, 18108354, 264288, 264688, 21906785, 21906766, 21906768, 21906789, 265022, 264652, 264683, 264628, 35695855, 264596, 264637, 264657, 264598, 83373044, 98264486, 22278900, 22278902, 264584
1652	86598622 (3303, 3304)	Novel Protein sim. GBank gi 1657437 (U73200) - p118Rip [Mus musculus]	Contains protein domain (PF00768) - situd PH domain	22278997, 28196498, 36162435, 21906754, 264638, 265018, 265019, 264909, 58844150, 33657109, 22278900, 22278902
1653	94255923 (3305, 3306)	Novel Protein sim. GBank gi 377654 emb CA406273 - (AJ045999) Tapsin [Gallus gallus]	Contains protein domain (PF00047) - immunoglobulin domain	18108388, 22278995, 22278998, 264259, 29331822, 29331824, 60432289, 29331828, 29331827, 29331830, 264909, 265006, 265009, 60432229, 60433356, 60433438, 21906754, 265017, 265019, 264448, 264683, 264288, 265021, 265022, 264682, 18108384, 65274791, 18108384, 60432113, 264597, 33657109, 264965
1654	79758471 (3307, 3308)		UNCLASSIFIED	

1655	06608346 (3309, 3310)	Novel Protein sim. GBank g 3356317 emb CAA72496 - (Y13053) seryl-RNA synthetase [Eca may]		synthase	52844507, 35696286, 22278988, 22278989, 28331824, 29331825, 29331826, 33565970, 33565971, 33565972, 2645017, 2645018, 2645019, 2645020, 2645021, 2645022, 33657108, 52844332, 264557, 56162323, 56526486, 60432113
1656	79962297 (3311, 3312)	Novel Protein sim. GBank g 1890141 db BAA18947 - (O83206) P24 protein [Mus musculus]		UNCLASSIFIED	29331822, 29331824, 29331825, 264563
1657	87771994 (3313, 3314)	Novel Protein sim. GBank g 4557645 ref NP_001524 IpiNRP - heterogenous nuclear ribonucleoprotein L	Contains protein domain (PF00076) - RNA recognition motif. (a.k.a. RRM, RED, or RNP domain)	- dna_jna_dmd	29331827, 265009, 21906766, 21906767, 265020, 265022, 33657108, 264638, 56526486, 264482
1658	87773778 (3315, 3316)	Novel Protein sim. GBank g 3870712 emb CAA87060 - (Z46937) similarity with ribosomal protein L21 [Caenorhabditis elegans]	Contains protein domain (PF00829) - Ribosomal protkaryotic L21 protein	UNCLASSIFIED	52846355, 35696286, 22278986, 22278987, 2278999, 2278999, 293259, 29331822, 29331824, 29331825, 29331826, 29331827, 29331828, 29166489, 264905, 264908, 52844042, 285006, 60433356, 264757, 60433438, 21906754, 265011, 18100351, 264448, 264385, 264288, 264768, 264798, 21906753, 21906767, 21906768, 21906769, 264863, 264711, 265022, 18100352, 264863, 33657108, 264863, 33657108, 264863, 33657108, 18100353, 22279000, 264482
1659	88230101 (3317, 3318)	Novel Protein sim. GBank g 533218 p S38018 - hypothetical protein YKL201c - yeast (Saccharomyces cerevisiae)		UNCLASSIFIED	52846317, 21906766, 21906767, 21906768, 87168516, 22278986, 265020, 22278989, 87168559, 264603, 265017, 264631, 265018, 265019, 22279002, 264482, 264635, 284565
1660	94315513 (3319, 3320)	Novel Protein sim. GBank g 2487012 p Q1001QYSA_CAAEEL - HYPOTHETICAL 26.8 KD PROTEIN THC3.4 IN CHROMOSOME III		UNCLASSIFIED	264488, 35696286, 264258, 33565952, 264508, 264509, 264485, 264896, 264907, 264908, 264909, 264510, 264511, 265006, 265007, 264512, 265009, 264810, 264582, 264596, 265010, 265009, 264602, 265017, 265018, 264503, 264765, 264784, 264888, 264786, 264686, 264768, 264789, 264689, 21906766, 33565987, 264690, 33657023, 264693, 33657108, 264689, 335659423, 33565953, 264634, 264603, 264555, 264638, 264639, 264640, 264641, 264642, 264643, 18100354, 18100355, 18100356, 18100357, 18100358, 264593, 264483, 264564, 264655, 264566, 264485, 264567

1661	94234071 (3321, 3322)	Novel Protein sim. GBank g1475910gtrpnp_004759.1gSFAS - splicing factor, arginine-aspartic-rich 11	Contains protein domain (PF00076) - RNA recognition motif, (a.k.a. RRM, RBD, or RNP domain)	UNCLASSIFIED	264488, 22278999, 264259, 29331824, 29331828, 29331827, 29331828, 264506, 6671502, 29331830, 284908, 52544045, 265007, 264512, 6043356, 60433438, 264512, 264512, 264512, 264448, 264768, 264768, 264768, 264768, 21905769, 265020, 33657023, 33657109, 65274791, 87168518, 264482, 264563, 264564, 264565, 264567
1662	94135172 (3323, 3324)	Novel Protein sim. GBank g1170502g1p52875jP27_MOUSE - TRANSMEMBRANE PROTEIN PF127			18108392, 29331822, 29331828, 20281100, 264106, 265006, 265007, 265008, 18108348, 21906766, 18108385, 18108366, 18108374, 83373044, 18108385
1663	94217146 (3325, 3326)	Novel Protein sim. GBank g1484136(emb)CA43275.1 - (A050107) hypothetical protein [Homo sapiens]	Contains protein domain (PF00397) - WW domain	kinase	52845156, 58182575, 22278994, 22278985, 35966286, 22278996, 56984075, 22278987, 22278994, 22278999, 264259, 29331822, 29331828, 29331827, 29331828, 33656970, 29331830, 264908, 56182435, 264511, 60433556, 33657402, 33108954, 87168474, 87168558, 265017, 265018, 264605, 18108392, 264512, 264512, 264512, 264768, 21905769, 21905765, 21905765, 21905768, 21905768, 265021, 265022, 264631, 33657023, 264693, 265387, 33657109, 264630, 52644332, 83373044, 87168518, 60432113, 22279000
1664	94234076 (3327, 3328)	Novel Protein sim. GBank g13043692jBjBAX25510j - (A0011156) KIAA00584 protein [Homo sapiens]		UNCLASSIFIED	264488, 263394, 35666286, 29331824, 35966052, 264508, 264509, 264605, 264906, 264907, 264908, 264909, 264510, 264511, 265009, 264910, 60170831, 264591, 264592, 264595, 87168474, 265011, 264600, 264601, 264604, 264605, 264760, 264762, 18108351, 264681, 264682, 264763, 264683, 264764, 264768, 264884, 264788, 264887, 264789, 264788, 21905764, 21905765, 21905767, 35959377, 265021, 264534, 60170615, 264690, 264691, 264692, 33657109, 33657112, 264693, 6036370, 264629, 33657121, 264694, 264695, 264696, 264555, 264636, 264637, 264638, 264639, 264558, 83373044, 87168518, 264563, 264566, 264488

1665	Q1226942 (3378, 3330)	Novel Protein sim. GBank. q1083306[pi][S50065 - salicoidesin - mouse]	Contains protein domain (PF00047) - immunoglobulin domain	264488, 26331856, 26331026, 2645639, 264908, 264508, 264403, 264510, 264511, 264910, 264552, 264593, 264593, 264758, 264596, 264650, 264780, 264785, 264786, 264766, 264758, 264629, 264630, 264634, 264636, 8337304, 264564, 264566, 264587, 264486
1666	96358160 (3331, 3332)	Novel Protein sim. GBank. q1091343[isp0439430X6 - SCAPO - PUTATIVE PRE-MRNA SPLICING FACTOR ATP-DEPENDENT RNA HELICASE SPACU16.02C]	Contains protein domain (PF00705) - helicase S1 RNA binding domain	56994075, 23278999, 264459, 293331824, 29331826, 29331827, 29164486, 265009, 33109954, 9768555, 265019, 264288, 264688, 21908767, 21908769, 264691, 33657182, 18108370, 18108374, 18108385, 22279002
1667	Q1228655 (3333, 3334)	Novel Protein sim. GBank. q1568393[spj0430351 - (A8028022) KIAA1099 protein [Homo sapiens]]	Contains protein domain (PF01412) - transport Putative GTP-ase activating protein for Arf	26331857, 26331026, 264695, 264908, 264510, 265009, 264495, 264758, 265011, 26188555, 265017, 265018, 265019, 264448, 264786, 264688, 21908765, 21908767, 21908769, 265020, 265021, 60170615, 264690, 264692, 264693, 18108388, 18108370, 263372, 55810784, 294555, 83373044, 60432113, 22279000, 22278902
1668	86095135 (3335, 3336)	Novel Protein sim. GBank. q1207689[lgaa453883 1] - (AF002127) short region of weak similarity to protein kinase C; contains similarity to Pfam domain PFK0130 (DAG_P1E_bind). Score=10.0, E-value=0.034, N=1 [Caenorhabditis elegans]	kinase	6671417, 264508, 264569, 264606, 264907, 264908, 264511, 264910, 264764, 264687, 264689, 33657109, 33696423, 33695955, 264632
1669	Q1227846 (3337, 3338)	Novel Protein sim. GBank. q1387531[icm1][CA5414 1] - (Z69646) contains a yelmine and arginine rich domain, possesses weak similarity with the RNA binding domain from RNA splicing factor U2AF 65 KD cDNA: cDNA EST EMBL:D64658 comes from this gene; cDNA EST EMBL:D66829 comes fr...	UNCLASSIFIED	29331855, 33108984, 264568, 264787, 264698, 33657108, 3373704
1670	87628009 (3339, 3340)		UNCLASSIFIED	264259, 29331824, 29331827, 60433438, 265022, 264636
1671	87346372 (3341, 3342)	Novel Protein sim. GBank. q169645[isp042449K01 - YEAST - PROBABLE SERINE/THREONINE/PROTEIN KINASE YKL101W]	kinase Contains protein domain (PF00049) Eukaryotic protein kinase domain	33569286, 22278997, 29331825, 264909, 21908754, 265017, 265018, 265019, 264682, 264683, 264758, 264686, 21908766, 21908767, 21908769, 21908769, 264691, 264555, 264696, 22779000, 264566, 264906, 264909, 264632, 18108381
1672	86291834 (3343, 3344)	Novel Protein sim. GBank. q11614270 (U74568) - double-stranded RNA specific adenosine deaminase [Rattus norvegicus]	deaminase	

1633	94318213 (3385, 3386)	Novel Protein sim. GBank g15033771p1p1MP_005704.1p1GPPNP - goodpasture antigen- binding protein	Contains protein domain (PF01655) - START domain	UNCLASSIFIED	263994, 35696296, 35696052, 264505, 264509, 264605, 264908, 264907, 264906, 264902, 264511, 265006, 265007, 265009, 265010, 264935, 264782, 264910, 265011, 264908, 264783, 264784, 264785, 264905, 264768, 264768, 35695917, 264692, 33657109, 35690453, 35690855, 264634, 264635, 264636, 264638, 264639, 33373044, 264486
1634	860054309 (3367, 3368)	Novel Protein sim. GBank g11255371 (U53147) - coded for by C. elegans cDNA yk34a9.5, coded for by C. elegans cDNA yk34a9.3. Similar to guanyle kinase. [Caenorhabditis elegans]	Contains protein domain (PF00623) - Guanyle kinase	UNCLASSIFIED	60424178, 52646842, 22278994, 35696296, 22278998, 264259, 52645080, 29331824, 29331828, 265007, 33657084, 265018, 264681, 264448, 264683, 264369, 264689, 21906765, 21906787, 21906764, 21906769, 265021, 264692, 65274620, 33657109, 27486282, 264635, 52644332, 56182233, 22279900
1635	87820710 (3371, 3372)	Novel Protein sim. GBank g12244707p1p1BAJ21115.11 - [A6052627] thrombospondin 1 [bos taurus]		UNCLASSIFIED	56182575, 264259, 264509, 264505, 264505, 264509, 264907, 264908, 264910, 264911, 265006, 265010, 264935, 264782, 264910, 265011, 264908, 264783, 264784, 264785, 265011, 264905, 18108361, 264764, 264765, 18108357, 264768, 18108362, 264628, 264630, 264631, 264634, 264635, 264637, 264638, 264639, 264565, 264486, 264567, 35695917, 264906, 264907, 264908, 264510, 265006, 265007, 264910, 264556, 18108381, 18108383, 265011
1637	94719400 (3373, 3374)	Novel Protein sim. GBank g14680079p1p1MA027729.1AF13295 - (AF132354) COL-20 protein [Homo sapiens]		UNCLASSIFIED	264905, 264910, 264760, 264629, 264555, 264569, 35696288, 22278999, 264508, 264908, 264905, 56182435, 33657402, 60433438, 55912038, 265017, 265018, 264448, 264764, 264288, 264686, 29146829, 35695917, 265020, 265021, 263972, 18108374, 85274797, 83373044, 264089
1638	82158442 (3375, 3376)	Novel Protein sim. GBank g14240193p1p1BAA74875.11 - [A9020659] KIAA0852 protein [Homo sapiens]		UNCLASSIFIED	264905, 264910, 264760, 264629, 264555, 264569, 35696288, 22278999, 264508, 264908, 264905, 56182435, 33657402, 60433438, 55912038, 265017, 265018, 264448, 264764, 264288, 264686, 29146829, 35695917, 265020, 265021, 263972, 18108374, 85274797, 83373044, 264089
1639	81255346 (3379, 3380)	Novel Protein sim. GBank g1360726 (AF031572) - seven- less-like protein kinase [Caenorhabditis elegans]	Contains protein domain (PF00028) - Cathelin domain	UNCLASSIFIED	264905, 264910, 264760, 264629, 264555, 264569, 35696288, 22278999, 264508, 264908, 264905, 56182435, 33657402, 60433438, 55912038, 265017, 265018, 264448, 264764, 264288, 264686, 29146829, 35695917, 265020, 265021, 263972, 18108374, 85274797, 83373044, 264089
1691	80095223 (3381, 3382)	Novel Protein sim. GBank g1360726 (AF031572) - seven- less-like protein kinase [Caenorhabditis elegans]	Contains protein domain (PF00028) - Cathelin domain	UNCLASSIFIED	264905, 264910, 264760, 264629, 264555, 264569, 35696288, 22278999, 264508, 264908, 264905, 56182435, 33657402, 60433438, 55912038, 265017, 265018, 264448, 264764, 264288, 264686, 29146829, 35695917, 265020, 265021, 263972, 18108374, 85274797, 83373044, 264089
1692	86106709 (3383, 3384)	Novel Protein sim. GBank g1121271p1p1p20207SLB_LAMEL - GLOBIN	Contains protein domain (PF00042) - Globin	UNCLASSIFIED	264106, 29331828, 264503, 264905, 264907, 264556, 265010, 265011, 21906768, 33657023, 264929, 263975, 264558

1694	94208168 (3397, 3398)	Novel Protein sim. GBank g16453932(e)P_008225, 1pPOLAR - polymerase (RNAs II (DNA directed) polypeptide J (13, 90D)	Combines protein domain (PF01193) RNA polymerase L / 13 to 19 kDa subunit	36552606, 22276996, 22276998, 22276999, 24533, 29331622, 29331625, 29331626, 29331628, 29331629, 29331630, 29331631, 29331632, 29331633, 29331634, 29331635, 29331636, 29331637, 29331638, 29331639, 29331640, 29331641, 29331642, 29331643, 29331644, 29331645, 29331646, 29331647, 29331648, 29331649, 29331650, 29331651, 29331652, 29331653, 29331654, 29331655, 29331656, 29331657, 29331658, 29331659, 29331660, 29331661, 29331662, 29331663, 29331664, 29331665, 29331666, 29331667, 29331668, 29331669, 29331670, 29331671, 29331672, 29331673, 29331674, 29331675, 29331676, 29331677, 29331678, 29331679, 29331680, 29331681, 29331682, 29331683, 29331684, 29331685, 29331686, 29331687, 29331688, 29331689, 29331690, 29331691, 29331692, 29331693, 29331694, 29331695, 29331696, 29331697, 29331698, 29331699, 29331700, 29331701, 29331702, 29331703, 29331704, 29331705, 29331706, 29331707, 29331708, 29331709, 29331710, 29331711, 29331712, 29331713, 29331714, 29331715, 29331716, 29331717, 29331718, 29331719, 29331720, 29331721, 29331722, 29331723, 29331724, 29331725, 29331726, 29331727, 29331728, 29331729, 29331730, 29331731, 29331732, 29331733, 29331734, 29331735, 29331736, 29331737, 29331738, 29331739, 29331740, 29331741, 29331742, 29331743, 29331744, 29331745, 29331746, 29331747, 29331748, 29331749, 29331750, 29331751, 29331752, 29331753, 29331754, 29331755, 29331756, 29331757, 29331758, 29331759, 29331760, 29331761, 29331762, 29331763, 29331764, 29331765, 29331766, 29331767, 29331768, 29331769, 29331770, 29331771, 29331772, 29331773, 29331774, 29331775, 29331776, 29331777, 29331778, 29331779, 29331780, 29331781, 29331782, 29331783, 29331784, 29331785, 29331786, 29331787, 29331788, 29331789, 29331790, 29331791, 29331792, 29331793, 29331794, 29331795, 29331796, 29331797, 29331798, 29331799, 29331800, 29331801, 29331802, 29331803, 29331804, 29331805, 29331806, 29331807, 29331808, 29331809, 29331810, 29331811, 29331812, 29331813, 29331814, 29331815, 29331816, 29331817, 29331818, 29331819, 29331820, 29331821, 29331822, 29331823, 29331824, 29331825, 29331826, 29331827, 29331828, 29331829, 29331830, 29331831, 29331832, 29331833, 29331834, 29331835, 29331836, 29331837, 29331838, 29331839, 29331840, 29331841, 29331842, 29331843, 29331844, 29331845, 29331846, 29331847, 29331848, 29331849, 29331850, 29331851, 29331852, 29331853, 29331854, 29331855, 29331856, 29331857, 29331858, 29331859, 29331860, 29331861, 29331862, 29331863, 29331864, 29331865, 29331866, 29331867, 29331868, 29331869, 29331870, 29331871, 29331872, 29331873, 29331874, 29331875, 29331876, 29331877, 29331878, 29331879, 29331880, 29331881, 29331882, 29331883, 29331884, 29331885, 29331886, 29331887, 29331888, 29331889, 29331890, 29331891, 29331892, 29331893, 29331894, 29331895, 29331896, 29331897, 29331898, 29331899, 29331900, 29331901, 29331902, 29331903, 29331904, 29331905, 29331906, 29331907, 29331908, 29331909, 29331910, 29331911, 29331912, 29331913, 29331914, 29331915, 29331916, 29331917, 29331918, 29331919, 29331920, 29331921, 29331922, 29331923, 29331924, 29331925, 29331926, 29331927, 29331928, 29331929, 29331930, 29331931, 29331932, 29331933, 29331934, 29331935, 29331936, 29331937, 29331938, 29331939, 29331940, 29331941, 29331942, 29331943, 29331944, 29331945, 29331946, 29331947, 29331948, 29331949, 29331950, 29331951, 29331952, 29331953, 29331954, 29331955, 29331956, 29331957, 29331958, 29331959, 29331960, 29331961, 29331962, 29331963, 29331964, 29331965, 29331966, 29331967, 29331968, 29331969, 29331970, 29331971, 29331972, 29331973, 29331974, 29331975, 29331976, 29331977, 29331978, 29331979, 29331980, 29331981, 29331982, 29331983, 29331984, 29331985, 29331986, 29331987, 29331988, 29331989, 29331990, 29331991, 29331992, 29331993, 29331994, 29331995, 29331996, 29331997, 29331998, 29331999, 29332000, 29332001, 29332002
------	-----------------------	---	---	---

1707	86041230 (3413, 3414)	Novel Protein sim. GBank gI4321664(gbI4AD1597) - (A7055470) ZNF228 [Homo sapiens]		UNCLASSIFIED	18108396, 22278997, 2642559, 29147620, 29331826, 29146498, 264905, 264906, 265008, 264493, 264695, 264738, 264986, 265018, 264780, 18108351, 264764, 264766, 264688, 264693, 18108370, 35690423, 264737, 264436, 87160510, 60432113, 264563
1708	91220519 (3415, 3416)	Novel Protein sim. GBank gI517459 (jptHP_005947, jptHTF - 510, methyltetrahydrofolate dehydrogenase, 5,10-methyltetrahydrofolate cyclohydrolase, 10-formylmethylenetetrahydrofolate synthetase	Contains protein domain (PF01269) - Formate-tetrahydrofolate ligase	synthase	56182576, 22278998, 5689404, 264259, 29331822, 26331824, 29331828, 29331827, 29331828, 29146498, 29146495, 29331830, 265009, 60170031, 33657402, 33109554, 87168559, 265019, 18108351, 284448, 21908765, 21906767, 21906766, 29186827, 29148629, 29148784, 60170815, 5264150, 33657023, 33657109, 18108374, 55011576, 264559, 18108385, 22279000, 264563
1709	90222583 (3417, 3418)			UNCLASSIFIED	264107, 55811957, 263974, 263976, 263977, 263981
1710	20754572 (3419, 3420)	Novel Protein sim. GBank gI501733(jptHP_005760, jpt IEC - N-acetylglucosamine		UNCLASSIFIED	55274572, 29331824, 29331826, 264768, 50431528, 35694423, 60432113, 264953
1711	91013729 (3421, 3422)	Novel Protein sim. GBank gI545416(jptHP_006463, jptAP4 - HBV associated factor.	Contains protein domain (PF00841) - Zn-finger in Ran binding protein and others.	kinase	56894075, 264693, 2642439, 29331822, 265009, 29331824, 29331827, 264107, 264110, 264611, 264332, 265011, 265018, 264683, 264696, 264689, 265020, 33657023, 263197, 33657109, 263974, 35696423, 35695855, 264630, 264636, 264558, 284566
1712	95330184 (3423, 3424)	Novel Protein sim. GBank gI150409 (M89183) - mature, parasite-infected erythrocyte surface antigen [Plasmodium falciparum]	Contains protein domain (PF00843) - B-box zinc finger.	UNCLASSIFIED	22278995, 264308, 264758, 18108351, 18108370, 263974, 18108374, 264634, 56182323, 83373044, 60432113
1713	64143453 (3425, 3426)	Novel Protein sim. GBank gI150409 (M89183) - mature, parasite-infected erythrocyte surface antigen [Plasmodium falciparum]		UNCLASSIFIED	22278997, 264757, 21906765, 265020, 265021, 264652, 5652466
1714	87420048 (3427, 3428)	Novel Protein sim. GBank gI5695337(M89183A03082.1) - (A802023) KIAA1100 protein [Homo sapiens]		UNCLASSIFIED	264909, 264902, 264906, 264907, 264908, 264909, 264910, 264901, 265011, 264766, 264768, 264769, 264681, 264682, 264632, 264634, 264635, 264636, 264637, 264656, 264639, 264654
1715	94280257 (3429, 3430)	Novel Protein sim. GBank gI4586168(jptEA7851.1) - (A802023) KIAA1100 protein [Homo sapiens]		transport	264639, 264654, 60432113, 60432116, 264109, 264909, 265007, 264600, 265019, 264686, 265020, 264693, 55811576, 264658, 60432113, 22279002
1716	87400449 (3431, 3432)	Novel Protein sim. GBank gI4586168(jptEA7851.1) - (A802023) KIAA1100 protein [Homo sapiens]			

1723	95358181 (3445, 3446) Novel Protein sim. GBank gll426562(gb AAD2063) - (AF-126062) Art-like 2 binding protein BART1 [Homo sapiens]	UNCLASSIFIED	264488, 264637, 264769, 21906767, 21906768, 56182575, 55811957, 22276897, 22276898, 265020, 264259, 264692, 33857023, 2933182, 29331824, 29331825, 60432286, 3357182, 3356970, 3357248, 29146496, 264695, 264607, 16103370, 264695, 264695, 264695, 264695, 55811576, 264510, 265008, 264511, 265007, 264910, 264632, 264691, 60432228, 264592, 60433356, 264594, 60433438, 264595, 83373044, 55812038, 33105954, 33057084, 87168518, 87168474, 265010, 265011, 87168559, 264500, 60432113, 264604, 265019, 264563, 264448, 264682, 264566, 264784, 264288, 264467, 264486, 264389, 264786
1724	87713856 (3447, 3448) Novel Protein sim. GBank gll2340162 (AF005083) - dsRBP- ZFs [Xenopus laevis]	UNCLASSIFIED	264905, 18100359, 264683, 264628, 264631, 264636, 264555, 264556, 264558, 264559
1725	86555151 (3449, 3450) Novel Protein sim. GBank gll3152652 (AF004904) - KE03 protein [Homo sapiens]	homeobox Contains protein domain (PF00023) Antk repeat	35896286, 264259, 29331822, 35896052, 264508, 264509, 264505, 264906, 264907, 264908, 264909, 264510, 265009, 264591, 264601, 264760, 18103351, 264681, 264764, 264682, 264683, 264684, 264685, 35895917, 264628, 35895423, 264630, 264631, 264632, 264635, 264636, 264638, 87168518, 264566
1726	85754255 (3451, 3452)	UNCLASSIFIED	29146488, 264483, 264689
1727	85298362 (3453, 3454) Novel Protein sim. GBank gll4689348(gb AAQ27261.1AF-13256) - (AF-132562) BcdNA LD14270 [Drosophila melanogaster]	UNCLASSIFIED	264905, 265011, 264689, 21906768

7128	95340515 (3465, 3466)	Novel Protein sm. GBank g 4405549(gBAA020027) - (AF131738) Unknown [Homo sapiens]	UNCLASSIFIED
1729	[1227948 (3457, 3458)]	Novel Protein sm. GBank g 854055(sCMA058337) - (S25413) U88 [Human testis/testis G]	UNCLASSIFIED
1730	85483474 (3458, 3459)	Novel Protein sm. GBank g 631600(pj S4-7094 - hypophthalic protein - rabbit	UNCLASSIFIED
1731	80265068 (3461, 3462)	Hypophthalic protein - rabbit	UNCLASSIFIED
1732	91218878 (3463, 3464)	Novel Protein sm. GBank g 4240231(BD BAAT4894.1) - (AB0000678) K/AA0871 protein [Homo sapiens]	struct
1733	81617178 (3465, 3466)	Novel Protein sm. GBank g 515755e (J70657.4) - m-Nub [Mus musculus]	Cysteine protein domain (PF00940) protein protease inhibition domain (PTPase) (PTBP/D).
1734	87856281 (3467, 3468)		

1735	86319538 (3469, 3470)	Novel Protein sim. GBank gl14833607/gp1A003056.1(AF-4679) PF727 [Mus musculus]	Contains protein domain (PF01169) - Uncharacterized protein family UPF0016	264488, 18108394, 18108398, 22278998, 60432049, 264259, 26331822, 264908, 265006, 265007, 265008, 264591, 33557402, 265010, 265011, 8716859, 264755, 264689, 21908765, 21908768, 264755, 264689, 21908765, 21908768, 21908769, 26148629, 26148784, 265021, 265022, 52644150, 18108364, 18108365, 33557108, 18108370, 18108374, 18108380, 18108385, 87168518, 264563, 18108390, 264490, 264259, 66714117, 66712502, 56182435, 265006, 265008, 264910, 60433356, 87168559, 265017, 265019, 18108351, 265020, 265022, 33557023, 33557108, 264555, 60431850, 264637, 60170394, 264559, 264639, 261964, 18108385, 87168518, 264591, 265011, 265019, 265015, 22279002
1736	95362884 (3471, 3472)	Novel Protein sim. GBank gl16885647/pef1NP_005472.1pTRAP - thyroid hormone receptor-associated protein complex component	UNCLASSIFIED	264490, 264259, 66714117, 66712502, 56182435, 265006, 265008, 264910, 60433356, 87168559, 265017, 265019, 18108351, 265020, 265022, 33557023, 33557108, 264555, 60431850, 264637, 60170394, 264559, 264639, 261964, 18108385, 87168518, 264591, 265011, 265019, 265015, 22279002
1737	88166549 (3473, 3474)	Novel Protein sim. GBank gl12143607/pf1S68695 - B/K protein - rat	Contains protein domain (PF00169) - C2 domain	264490, 264259, 66714117, 66712502, 56182435, 265006, 265008, 264910, 60433356, 87168559, 265017, 265019, 18108351, 265020, 265022, 33557023, 33557108, 264555, 60431850, 264637, 60170394, 264559, 264639, 261964, 18108385, 87168518, 264591, 265011, 265019, 265015, 22279002
1738	85768811 (3475, 3476)	Novel Protein sim. GBank gl12225941/emo1CA469714 - [Y08460] Males protein [Mus musculus]	UNCLASSIFIED	264490, 264259, 66714117, 66712502, 56182435, 265006, 265008, 264910, 60433356, 87168559, 265017, 265019, 18108351, 265020, 265022, 33557023, 33557108, 264555, 60431850, 264637, 60170394, 264559, 264639, 261964, 18108385, 87168518, 264591, 265011, 265019, 265015, 22279002
1739	87328576 (3477, 3478)	Novel Protein sim. GBank gl14833607/gp1A003056.1(AF-4679) PF727 [Mus musculus]	Contains protein domain (PF00441) - ACh-Cox GTPase	264490, 264259, 66714117, 66712502, 56182435, 265006, 265008, 264910, 60433356, 87168559, 265017, 265019, 18108351, 265020, 265022, 33557023, 33557108, 264555, 60431850, 264637, 60170394, 264559, 264639, 261964, 18108385, 87168518, 264591, 265011, 265019, 265015, 22279002
1740	83362939 (3479, 3480)	Novel Protein sim. GBank gl14833607/gp1A003056.1(AF-4679) PF727 [Mus musculus]	Contains protein domain (PF00441) - ACh-Cox GTPase	264490, 264259, 66714117, 66712502, 56182435, 265006, 265008, 264910, 60433356, 87168559, 265017, 265019, 18108351, 265020, 265022, 33557023, 33557108, 264555, 60431850, 264637, 60170394, 264559, 264639, 261964, 18108385, 87168518, 264591, 265011, 265019, 265015, 22279002
1741	95010100 (3481, 3482)	Novel Protein sim. GBank gl14833607/gp1A003056.1(AF-4679) PF727 [Mus musculus]	Contains protein domain (PF00383) - FIVE zinc finger	264490, 264259, 66714117, 66712502, 56182435, 265006, 265008, 264910, 60433356, 87168559, 265017, 265019, 18108351, 265020, 265022, 33557023, 33557108, 264555, 60431850, 264637, 60170394, 264559, 264639, 261964, 18108385, 87168518, 264591, 265011, 265019, 265015, 22279002
1742	85768814 (3483, 3484)	Novel Protein sim. GBank gl14505193/pef1NP_003667.1pMLD1 - membrane fatty acid lipid desaturase	UNCLASSIFIED	264490, 264259, 66714117, 66712502, 56182435, 265006, 265008, 264910, 60433356, 87168559, 265017, 265019, 18108351, 265020, 265022, 33557023, 33557108, 264555, 60431850, 264637, 60170394, 264559, 264639, 261964, 18108385, 87168518, 264591, 265011, 265019, 265015, 22279002
1743	86966475 (3485, 3486)	Novel Protein sim. GBank gl128832/pep19189ALU2 - HUMAN - III ALU SUBFAMILY SE WARNING ENTRY III	UNCLASSIFIED	264490, 264259, 66714117, 66712502, 56182435, 265006, 265008, 264910, 60433356, 87168559, 265017, 265019, 18108351, 265020, 265022, 33557023, 33557108, 264555, 60431850, 264637, 60170394, 264559, 264639, 261964, 18108385, 87168518, 264591, 265011, 265019, 265015, 22279002
1744	91224003 (3487, 3488)	Novel Protein sim. GBank gl128832/pep19189ALU2 - HUMAN - III ALU SUBFAMILY SE WARNING ENTRY III	UNCLASSIFIED	264490, 264259, 66714117, 66712502, 56182435, 265006, 265008, 264910, 60433356, 87168559, 265017, 265019, 18108351, 265020, 265022, 33557023, 33557108, 264555, 60431850, 264637, 60170394, 264559, 264639, 261964, 18108385, 87168518, 264591, 265011, 265019, 265015, 22279002
1745	20295075 (3489, 3490)	Novel Protein sim. GBank gl1317506/p3873HY16 - YEAST - HYPOPHARYNGEAL REGION	Contains protein domain (PF00440) - WD domain, G-beta repeat	264490, 264259, 66714117, 66712502, 56182435, 265006, 265008, 264910, 60433356, 87168559, 265017, 265019, 18108351, 265020, 265022, 33557023, 33557108, 264555, 60431850, 264637, 60170394, 264559, 264639, 261964, 18108385, 87168518, 264591, 265011, 265019, 265015, 22279002
1746	94326110 (3491, 3492)	Novel Protein sim. GBank gl1317506/p3873HY16 - YEAST - HYPOPHARYNGEAL REGION	Contains protein domain (PF00440) - WD domain, G-beta repeat	264490, 264259, 66714117, 66712502, 56182435, 265006, 265008, 264910, 60433356, 87168559, 265017, 265019, 18108351, 265020, 265022, 33557023, 33557108, 264555, 60431850, 264637, 60170394, 264559, 264639, 261964, 18108385, 87168518, 264591, 265011, 265019, 265015, 22279002
1747	94324333 (3493, 3494)	Novel Protein sim. GBank gl1317506/p3873HY16 - YEAST - HYPOPHARYNGEAL REGION	Contains protein domain (PF00440) - WD domain, G-beta repeat	264490, 264259, 66714117, 66712502, 56182435, 265006, 265008, 264910, 60433356, 87168559, 265017, 265019, 18108351, 265020, 265022, 33557023, 33557108, 264555, 60431850, 264637, 60170394, 264559, 264639, 261964, 18108385, 87168518, 264591, 265011, 265019, 265015, 22279002

1748	88005580 (2455, 3498)	Novel Protein sm. GBank glt150451 [lncfNP_001530, lprKSJ2 - heat shock protein, DNA-like 2]	Contains protein domain (PF00884) Dual central domain (4 repeats)	264480, 55182975, 2503192, 56182405, 264112, 26502975, 2503191, 21908768, 265020, 264691, 55815576, 264653, 264555, 254556, 264557, 264559	264106	Contains protein domain (PF00815) Regulator of G protein signaling domain	33657402, 264288, 32644150, 263374, 83373044	
1749	83353091 (2457, 3498)	Novel Protein sm. GBank glt6560780 [pAAO3948, 1AF15195 - (AF15196) RGS protein RGS-17 (Callis gallus)]	Contains protein domain (PF00849) Regulator of G protein signaling domain	264112, 26502975, 2503192, 56182405, 264692, 264784, 264539, 264288, 264685, 264687, 21906766, 264692, 264693, 65274620, 65274791, 31665555, 264637, 264564	264106	Contains protein domain (PF00849) Regulator of G protein signaling domain	265010, 254369	
1750	94321684 (2499, 3500)	Novel Protein sm. GBank glt2760161 [pAAO3948, 1AF15195 - (AF15196) RGS protein RGS-17 (Callis gallus)]	Contains protein domain (PF00849) Regulator of G protein signaling domain	264112, 26502975, 2503192, 56182405, 264692, 264784, 264539, 264288, 264685, 264687, 21906766, 264692, 264693, 65274620, 65274791, 31665555, 264637, 264564	264106	Contains protein domain (PF00849) Regulator of G protein signaling domain	265010, 254369	
1751	83373038 (3501, 3502)	Novel Protein sm. GBank glt2760161 [pAAO3948, 1AF15195 - (AF15196) RGS protein RGS-17 (Callis gallus)]	Contains protein domain (PF00849) Regulator of G protein signaling domain	264112, 26502975, 2503192, 56182405, 264692, 264784, 264539, 264288, 264685, 264687, 21906766, 264692, 264693, 65274620, 65274791, 31665555, 264637, 264564	264106	Contains protein domain (PF00849) Regulator of G protein signaling domain	265010, 254369	
1752	86465530 (3503, 3504)	Novel Protein sm. GBank glt91544236 [p7A4M4YK28, SNNY3 - HYPOTHETICAL 35.0 KD PROTEIN, SLR1023]	Contains protein domain (PF00849) Regulator of G protein signaling domain	264112, 26502975, 2503192, 56182405, 264692, 264784, 264539, 264288, 264685, 264687, 21906766, 264692, 264693, 65274620, 65274791, 31665555, 264637, 264564	264106	Contains protein domain (PF00849) Regulator of G protein signaling domain	265010, 254369	
1753	94235159 (3505, 3506)	Novel Protein sm. GBank glt2856359 (AF-007155) - unknown (Homo sapiens)	Contains protein domain (PF00849) Regulator of G protein signaling domain	264112, 26502975, 2503192, 56182405, 264692, 264784, 264539, 264288, 264685, 264687, 21906766, 264692, 264693, 65274620, 65274791, 31665555, 264637, 264564	264106	Contains protein domain (PF00849) Regulator of G protein signaling domain	265010, 254369	
1754	88095323 (3507, 3508)	Novel Protein sm. GBank glt731421 [pY39981 [YE4H, YEAST - HYPOTHETICAL 53 KD PROTEIN IN HXT6-CAN1 INTERGENIC REGION]	Contains protein domain (PF00849) Regulator of G protein signaling domain	264112, 26502975, 2503192, 56182405, 264692, 264784, 264539, 264288, 264685, 264687, 21906766, 264692, 264693, 65274620, 65274791, 31665555, 264637, 264564	264106	Contains protein domain (PF00849) Regulator of G protein signaling domain	265010, 254369	
1755	79470282 (3509, 3510)	Novel Protein sm. GBank glt1176422 (U43194) - Rnuplin (Mus musculus)	Contains protein domain (PF00849) Regulator of G protein signaling domain	264112, 26502975, 2503192, 56182405, 264692, 264784, 264539, 264288, 264685, 264687, 21906766, 264692, 264693, 65274620, 65274791, 31665555, 264637, 264564	264106	Contains protein domain (PF00849) Regulator of G protein signaling domain	265010, 254369	
1756	92662614 (3511, 3512)	Novel Protein sm. GBank glt432850 [p19A-D20708] - (A206300) putative glucose-induced repressor protein (Arabidopsis thaliana)	Contains protein domain (PF00849) Regulator of G protein signaling domain	264112, 26502975, 2503192, 56182405, 264692, 264784, 264539, 264288, 264685, 264687, 21906766, 264692, 264693, 65274620, 65274791, 31665555, 264637, 264564	264106	Contains protein domain (PF00849) Regulator of G protein signaling domain	265010, 254369	
1757	95357380 (3513, 3514)	Novel Protein sm. GBank glt541451 [p19A-D20708] - (A206300) putative glucose-induced repressor protein (Arabidopsis thaliana)	Contains protein domain (PF00849) Regulator of G protein signaling domain	264112, 26502975, 2503192, 56182405, 264692, 264784, 264539, 264288, 264685, 264687, 21906766, 264692, 264693, 65274620, 65274791, 31665555, 264637, 264564	264106	Contains protein domain (PF00849) Regulator of G protein signaling domain	265010, 254369	

1758	07612371 (3515, 3516)	Novel Protein sim. GBank gl3881040mjcAA16403 - (AL021497) predicted using GeneFinder (Caenorhabditis elegans)	UNCLASSIFIED	22278994, 22278999, 29331822, 29331824, 29331825, 33569970, 264908, 265006, 265007, 265009, 264911, 33569971, 33569972, 33569973, 33569974, 33569975, 265017, 265018, 21906769, 264920, 265021, 3357023, 3357109, 264629, 18108374, 35695955, 264632, 52644332, 22279002, 264653
1759	36994372 (3517, 3518)	Novel Protein sim. GBank gl52742748jmgCAB45688.11- (AJ13120)	UNCLASSIFIED	264759
1760	87326716 (3519, 3520)	Novel Protein sim. GBank gl52742748jmgCAB45688.11- (AJ13120) Proteins rich synapse associated protein 2 [Rattus norvegicus]	UNCLASSIFIED	264906, 264907, 264908, 264909, 265006, 265009, 264910, 60432229, 264562, 264565, 55812038, 264788, 264762, 18108351, 264764, 264766, 264768, 264769, 21906765, 55811957, 35699917, 264690, 264692, 264628, 264629, 5811576, 35699423, 264632, 264634, 264636, 264557, 264639, 60432113, 264595, 264466
1761	87405586 (3521, 3522)	Novel Protein sim. GBank gl1277493ppj059jMTSC_A0CA - MYOSIN I C-HEAVY CHAIN	UNCLASSIFIED	29331822, 264910, 264905, 264686
1762	95319867 (3523, 3524)	Novel Protein sim. GBank gl3189158 (AC004770) - BC268730.2 [Homo sapiens]	deaturate	22278999, 264259, 264605, 264907, 60170831, 265010, 265011, 265017, 264448, 21906765, 21906766, 21906767, 21906768, 265021, 264690, 33567109, 18108374, 264558, 60170394
1763	91224013 (3525, 3526)	Novel Protein sim. GBank gl4809026gpaAD30062.11 - (AF132856) suppressor of G2 allele of skp1 homolog [Homo sapiens]		56181580, 29331825, 35696095, 264805, 264906, 264909, 264763, 264682, 264769, 35695917, 265022, 33567023, 18108374, 35695423, 264634
1764	87757697 (3527, 3528)	Novel Protein sim. GBank gl1350669ppjICG-HU1V - collagen alpha 1(V) chain precursor - human	collagen	22278999, 264259, 35696095, 29331828, 356182435, 265008, 265017, 265018, 264448, 264288, 21906766, 21906767, 29148627, 35695917, 264691, 33567023, 60432113, 35695423, 264634
1765	91230091 (3529, 3530)	Novel Protein sim. GBank gl468506ppj353503 - finger protein neutralized - fruit fly (Drosophila melanogaster)	UNCLASSIFIED	22278999, 264259, 35696095, 22278997, 264906, 264907, 264908, 264909, 264910, 264911, 265007, 265008, 264910, 265009, 21906764, 265010, 265011, 264448, 265018, 265019, 18108351, 264682, 264448, 264369, 264428, 264766, 18108359, 21906766, 21906767, 29148627, 29148629, 35695917, 265020, 265021, 264692, 284628, 264629, 18108374, 263976, 264636, 83373044, 22279000, 22279002

1766	95061201 (3531, 3532)	Novel Protein sim. Gbank gi1391419 p3655 OGT_RAT - UDP-N- ACETYLGLUCOSAMINE-6-PHOSPHATE TRANSFERASE (UGT) PRECURSOR (UJAS)			glycoprotein	52645156, 87160559, 90170915, 33657023, 264693, 33057169, 27482651, 264555, 83370044
1767	87755988 (3533, 3534)	Novel Protein sim. Gbank gi1417843 p3655 CAI18263.1 - PRECURSOR (UJAS)			UNCLASSIFIED	265017, 265019, 264686, 264766, 265020, 264692
1768	8025216 (3535, 3536)	Novel Protein sim. Gbank gi11042K10.4 (novel protein) (homo sapiens)			UNCLASSIFIED	264563
1769	87388888 (3537, 3538)	Novel Protein sim. Gbank gi115204 p100796C1R_HUMAN - COMPLEMENT C1R COMPONENT PRECURSOR	Contains protein domain (PF00089) - Trypsin		UNCLASSIFIED complement	264488, 264768, 264789, 56182575, 55811957, 264690, 264991, 35096052, 264903, 264509, 264906, 264907, 264628, 264908, 264909, 264910, 264634, 264635, 264638, 264556, 264737, 264738, 55812038, 63274444, 264760, 264563, 264762, 264764, 264684, 264765
1770	95413144 (3539, 3540)	Novel Protein sim. Gbank gi115204 p100796C1R_HUMAN - COMPLEMENT C1R COMPONENT PRECURSOR	Contains protein domain (PF00089) - Trypsin		UNCLASSIFIED	264488, 264768, 264789, 56182575, 55811957, 264690, 264991, 35096052, 264903, 264509, 264906, 264907, 264628, 264908, 264909, 264910, 264634, 264635, 264638, 264556, 264737, 264738, 55812038, 63274444, 264760, 264563, 264762, 264764, 264684, 264765
1771	94233542 (3541, 3542)	Novel Protein sim. Gbank gi1391419 p3655 OGT_RAT - UDP-N- ACETYLGLUCOSAMINE-6-PHOSPHATE TRANSFERASE (UGT) SUBUNIT (OJGLQMC) TRANSFERASE P110 SUBUNIT)	Contains protein domain (PF00089) - Trypsin		UNCLASSIFIED	264488, 264768, 264789, 56182575, 55811957, 264690, 264991, 35096052, 264903, 264509, 264906, 264907, 264628, 264908, 264909, 264910, 264634, 264635, 264638, 264556, 264737, 264738, 55812038, 63274444, 264760, 264563, 264762, 264764, 264684, 264765
1772	87643510 (3543, 3544)	Novel Protein sim. Gbank gi1095844 p3655 MD134351.1 AF12136 - (AF121360) DNDHHC/NEW1 zinc finger protein 11 (Drosophila melanogaster)	Contains protein domain (PF00089) - Trypsin		UNCLASSIFIED	264488, 264768, 264789, 56182575, 55811957, 264690, 264991, 35096052, 264903, 264509, 264906, 264907, 264628, 264908, 264909, 264910, 264634, 264635, 264638, 264556, 264737, 264738, 55812038, 63274444, 264760, 264563, 264762, 264764, 264684, 264765
1773	94116824 (3545, 3546)	Novel Protein sim. Gbank gi1391419 p3655 OGT_RAT - UDP-N- ACETYLGLUCOSAMINE-6-PHOSPHATE TRANSFERASE (UGT) SUBUNIT (OJGLQMC) TRANSFERASE P110 SUBUNIT)	Contains protein domain (PF00089) - Trypsin		UNCLASSIFIED	264488, 264768, 264789, 56182575, 55811957, 264690, 264991, 35096052, 264903, 264509, 264906, 264907, 264628, 264908, 264909, 264910, 264634, 264635, 264638, 264556, 264737, 264738, 55812038, 63274444, 264760, 264563, 264762, 264764, 264684, 264765
1774	94232573 (3547, 3548)	Novel Protein sim. Gbank gi12456599 p1019031Y032_HUMAN - HYPOTHETICAL PROTEIN KIAA0032	Contains protein domain (PF00089) - Trypsin		UNCLASSIFIED	264488, 264768, 264789, 56182575, 55811957, 264690, 264991, 35096052, 264903, 264509, 264906, 264907, 264628, 264908, 264909, 264910, 264634, 264635, 264638, 264556, 264737, 264738, 55812038, 63274444, 264760, 264563, 264762, 264764, 264684, 264765

1775	95358330 (3549, 3550)	Novel Protein sim. GBank gl(458199)pbj(BAA708487) - (C00928) The KIAA0138 gene product is novel. [Homo sapiens]	UNCLASSIFIED	65274372, 56182575, 56584075, 35658286, 22278997, 29331822, 25331826, 60432289, 25331826, 35698032, 25331830, 60712602, 254628, 814535, 264511, 253078, 254690, 254628, 814535, 264511, 253078, 254690, 21908764, 85358542, 87188555, 264601, 265017, 265018, 265019, 264752, 264448, 264288, 264689, 21906765, 21906766, 21906767, 21906768, 265020, 265021, 265022, 264691, 33657023, 264692, 33657105, 27485281, 33657345, 18108370, 18108377, 35698423, 55811678, 35698585, 264632, 264634, 264636, 264639, 56182323, 83373044, 56526486, 87188518, 60432113, 22279000, 22279002, 264482, 264485, 264910
1776	94133758 (3551, 3552)	Novel Protein sim. GBank gl(4589676)pbj(BAA78857.1) - (A902320) KIAA1013 protein [Homo sapiens]	nuc_ept	56584075, 20331826, 265008, 37168474, 265008, 37168474, 265008, 37168474, 265008, 37168474, 21906765, 21906766, 21906767, 21906768, 265020, 264691, 33657023, 264692, 33657105, 27485281, 33657345, 18108370, 18108377, 35698423, 55811678, 35698585, 264632, 264634, 264636, 264639, 56182323, 83373044, 56526486, 87188518, 60432113, 22279000, 22279002, 264482, 264485, 264910
1777	87447171 (3553, 3554)	Novel Protein sim. GBank gl(219359)pbj(115)YOKI_SCHPO - HYPOHETICAL 116.5 KD PROTEIN C20638.09C IN CHROMOSOME 1		56584075, 20331826, 265008, 37168474, 265008, 37168474, 265008, 37168474, 265008, 37168474, 21906765, 21906766, 21906767, 21906768, 265020, 264691, 33657023, 264692, 33657105, 27485281, 33657345, 18108370, 18108377, 35698423, 55811678, 35698585, 264632, 264634, 264636, 264639, 56182323, 83373044, 56526486, 87188518, 60432113, 22279000, 22279002, 264482, 264485, 264910
1778	94851624 (3555, 3556)	Novel Protein sim. GBank gl(387564)pbj(CAA91454.1) - (Z65681) Similarity to Human rab13 protein (P1R Acc. No. A49647). Contains the ATP/GTP-binding site motif (PROSITE PS00071); cDNA EST ENBL1M69412 comes from this gene; cDNA EST yk212g 3 comes from this gene; cDNA EST yk212g.	UNCLASSIFIED	29331826, 25331827, 35698055, 264412, 265007, 265009, 265017, 265019, 264762, 18108351, 264769, 21906765, 21906766, 21906767, 21906768, 21906769, 265020, 265021, 264691, 264693, 18108370, 264556, 83373044
1779	94133758 (3557, 3558)	Novel Protein sim. GBank gl(4589676)pbj(BAA78857.1) - (A902320) KIAA1013 protein [Homo sapiens]	UNCLASSIFIED	29148827, 35698286, 26147620, 265006, 265007, 265008, 18108385, 85274727, 264482, 264589, 264765
1780	87023497 (3559, 3560)		Contains protein domain (PF00907) - Apolactin	264107, 33657109, 56525486
1781	80474727 (3561, 3562)	Novel Protein sim. GBank gl(75225)pbj(41237)C1XN_0A1 - CORTEXIN	UNCLASSIFIED	264568, 264905, 264639
1782	86094607 (3563, 3564)			264495, 264589, 264639, 264690, 264695, 264910, 264903, 264208, 265007, 265009, 264910, 264991, 264745, 264764, 264288, 264768, 264769, 264635, 264636, 264637, 264639, 264583
1783	85717605 (3565, 3566)	Novel Protein sim. GBank gl(2257443)pbj(A21436) - (A9004538) protein arginine N methyltransferase [Schistosoma haematophylos pombe]	interferon	264768

1784	95197093 (3567, 3568)	Novel Protein sim. GBank gi175049 (U5542) - myosin X [Bos taurus]	Centelin protein domain (PF007169) - strand PH domain	25695265, 264259, 33690052, 364508, 264905, 264906, 264907, 66715502, 264908, 264909, 265007, 265008, 265009, 264910, 264591, 264594, 264757, 264758, 264759, 265010, 265011, 264801, 264602, 264904, 264605, 18103351, 264762, 264763, 264764, 264389, 264766, 264687, 264768, 264688, 21900768, 35605917, 254690, 264691, 264692, 264693, 264628, 18103374, 35695423, 294031, 264632, 264635, 264637, 264638, 264635, 18103365, 22279000, 22279002, 264595, 264596, 264446
1785	95357475 (3569, 3570)	Novel Protein sim. GBank gi459532db [BLAAT788.1] - (AB023171) KIAA0954 protein [Homo sapiens]	UNCLASSIFIED	95274572, 264436, 293316, 2535824, 264595, 264596, 264597, 264598, 264905, 264906, 66712362, 56182435, 264511, 265007, 60433366, 59811150, 264683, 264365, 264687, 52644229, 21900757, 52644150, 33657023, 65274620, 33657185, 65274791, 35695855, 264555, 65274727, 22279002
1786	85286465 (3571, 3572)	Novel Protein sim. GBank gi117783 [gip2870CYAA_RAT ADENYLATE CYCLASE, TYPE IV (ATP PYROPHOSPHATE-LYASE) (ADENYLATE CYCLASE)]	UNCLASSIFIED	264906, 35696423, 264636
1787	87434784 (3573, 3574)	Novel Protein sim. GBank gi387717 [emb] CAA90335.1] - (Z50028) cDNA EST x32108.5 comes from this gene; cDNA EST EMBL D68806 comes from this gene; cDNA EST x33956.5 comes from this gene [Caenorhabditis elegans]	UNCLASSIFIED	264488, 264905, 264906, 264909, 264595, 264764, 264765, 264692, 60431528, 264629, 264636, 264561, 264566
1788	91228779 (3575, 3576)	Novel Protein sim. GBank gi2088666 (AF003130) - F55A12.2 gene product [Caenorhabditis elegans]	UNCLASSIFIED	264488, 8337204
1789	8604529 (3577, 3578)	Novel Protein sim. GBank gi2088666 (AF003130) - F55A12.2 gene product [Caenorhabditis elegans]	UNCLASSIFIED	264488, 265021, 264555, 264639
1790	8248734 (3579, 3580)			35695052, 264905, 264906, 264907, 264908, 264699, 265008, 264910, 264758, 265011, 265019, 264764, 264765, 264769, 264628, 234635

1797	91221219 (3593, 3594)	Novel Protein sim. GBank g1154211 (U87156) - decoy (Arabidopsis thaliana)	inosomiprot	22278996, 22278997, 22278998, 22278999, 26331822, 264910, 60170031, 21906754, 52644225, 21906765, 21906768, 21906769, 35959317, 253022, 53544150, 264691, 33557023, 253097, 35957109, 22279000, 22279001, 22279002, 22279003, 22279004, 22279005, 22279006, 22279007, 22279008, 22279009, 264459, 26331826, 60422226, 29331828, 33656970, 265008, 60432228, 264757, 60433438, 21906754, 33657094, 87168559, 255017, 18108351, 264682, 264448, 264288, 21906765, 21906766, 21908767, 21906768, 21906769, 35695817, 265020, 265021, 33657023, 33657182, 27466261, 27466265, 33657349, 263973, 18108374, 55811576, 35695855, 18108385, 87168518, 22279000, 264448
1798	91221216 (3595, 3596)	Novel Protein sim. GBank g1263000 (p1A24608.1) - (D98340) dipeptidyl peptidase III (Rattus norvegicus)	peptidase	22278996, 22278997, 22278998, 22278999, 26331822, 264910, 60170031, 21906754, 52644225, 21906765, 21906768, 21906769, 35959317, 253022, 53544150, 264691, 33557023, 253097, 35957109, 22279000, 22279001, 22279002, 22279003, 22279004, 22279005, 22279006, 22279007, 22279008, 22279009, 264459, 26331826, 60422226, 29331828, 33656970, 265008, 60432228, 264757, 60433438, 21906754, 33657094, 87168559, 255017, 18108351, 264682, 264448, 264288, 21906765, 21906766, 21908767, 21906768, 21906769, 35695817, 265020, 265021, 33657023, 33657182, 27466261, 27466265, 33657349, 263973, 18108374, 55811576, 35695855, 18108385, 87168518, 22279000, 264448
1799	86321713 (3597, 3598)	Novel Protein sim. GBank g1658554 (p1A45054.1) - (A0029025) KIAA1102 protein (Homo sapiens)	eph	264908, 21906754, 21906767, 21906769, 265020, 33657023, 264692, 264693, 264404, 22279000
1800	87080116 (3599, 3600)	Novel Protein sim. GBank g14680079 (p1A027729.1) (AF132954) CGI-20 protein (Homo sapiens)	UNCLASSIFIED	264691, 264538, 264586
1801	95060723 (3601, 3602)	Novel Protein sim. GBank g1134920 (p1Z1997) (SSGF VOLCA - SULFATED SURFACE GLYCOPROTEIN 185 (SSG 185))		52584045, 265007, 264632
1802	87771012 (3603, 3604)	Novel Protein sim. GBank g1134920 (p1Z1997) (SSGF VOLCA - SULFATED SURFACE GLYCOPROTEIN 185 (SSG 185))		335956286, 65714117, 264508, 264509, 56162435, 264512, 18106351, 264688, 55811957, 264692, 55811576, 35695855, 264448
1803	95060725 (3605, 3606)	Novel Protein sim. GBank g14680079 (p1A027729.1) (AF132954) CGI-20 protein (Homo sapiens)		264696, 264488, 264687, 264689, 264768, 35959286, 35959317, 264259, 264691, 264692, 264693, 2261009, 18106354, 35958052, 264508, 264509, 264305, 264906, 18106370, 264628, 264907, 65712502, 264908, 264909, 18106374, 18106376, 35959423, 35959555, 264510, 264511, 359506, 265007, 264512, 265008, 264510, 264631, 264632, 264633, 264635, 264591, 264638, 264637, 264692, 264768, 264592, 18106353, 18106359, 264694, 53373044, 265011, 264760, 264683, 18106351, 18106352, 264584, 264585, 264763, 264683, 264764, 264448, 264565, 264765, 264693, 264766, 264595, 264768, 264486, 264567, 264765, 264766

1804	87770203 (5807, 3608)	Novel Protein sim. GBank g13879914[emc][CAA9538.1]- (ZNF43) predicted using Genefinder; cDNA EST EMBL C13890 comes from this gene; cDNA EST ENBL C11575 comes from this gene; cDNA EST yA3K4.5 comes from this gene [Caenorhabditis elegans]			52648385, 22278997, 22278998, 264985, 264986, 264989, 264910, 21800754, 264766, 21900765, 21900788, 33659171, 265520, 265521, 264637, 264638, 22279000, 264684, 264686
1805	95320375 (5608, 3610)	Novel Protein sim. GBank g1455644[hipp_008461.1] (pEBBP - estrogen-responsive B box protein)			29331824, 29331825, 29331826, 29331827, 29331828, 87169559, 264636, 6043213, 52540429, 36986423, 264636, 6043213
1806	94133762 (3611, 3612)	Novel Protein sim. GBank g14585676[huBAX76857.1]- (AB023220) KIAA1013 protein [Homo sapiens]		struct	264004, 264105, 264608, 35994423, 265008, 265007, 265008, 264555, 264592, 265011, 265018, 264569
1807	86943032 (3613, 3614)				28331824, 264908, 264910, 33657023, 263978
1808	87642711 (3615, 3616)	Novel Protein sim. GBank g14884079[emc][CAB43235.1]- (AL050008) hypothetical protein [Homo sapiens]		UNCLASSIFIED	264488, 35658286, 66714117, 35696052, 86712502, 264932, 60434338, 52644296, 265010, 264683, 264369, 264889, 55811957, 35695917, 33657109, 35695763, 55810764, 18106379, 35696423, 35695955, 5616232, 264636, 264637, 264638, 264639
1809	95321488 (3617, 3618)	Novel Protein sim. GBank g19191627 (u97965) - putative G protein [Mus musculus]		UNCLASSIFIED	29331824, 264908, 264620, 18100374, 264556, 264657, 264658
1810	88096316 (3619, 3620)	Novel Protein sim. GBank g113329446[p7178Y]pP. YEAST - HYPOTHETICAL 118.4 KD PROTEIN IN BAT2-DALS INTERGENIC REGION PRECURSOR		UNCLASSIFIED	264488, 35696052, 264905, 264906, 264907, 264908, 264909, 264511, 265009, 264910, 264592, 264593, 264594, 33657402, 264757, 264595, 264758, 264596, 264759, 264600, 264601, 264762, 264683, 264764, 264288, 264684, 264766, 264767, 264686, 264768, 264687, 264769, 264689, 265021, 264690, 264691, 264693, 264628, 264659, 18100374, 264630, 264631, 264632, 264634, 264635, 264636, 264637, 264638, 264639, 264693, 264556, 264486, 264697
1811	85088272 (3621, 3622)	Novel Protein sim. GBank g12134864[pir13725 - death- associated protein kinase (EC 2.7.1.3) - human]		Contains protein domain (PF-00023) - kinase Atk repeat	264908, 264909, 264511, 265009, 264910, 264592, 264593, 264594, 33657402, 264757, 264595, 264758, 264596, 264759, 264600, 18100374, 264766, 264686, 264769, 264534, 80170815, 33657023, 264629, 264631, 264639, 264693, 264682, 264483
1812	78245772 (3623, 3624)				29331822, 29331824, 265019, 18100351, 21905769

1813	88055972 (3625, 3626)	Novel Protein sim. GBank g1505 (638)g1AD38325.1AF07372 - (AF073727) EH domain-binding mitotic phosphoprotein [Homo sapiens]	Contains protein domain (PF01417) ENTH domain	glucosylase	56182575, 264259, 29331824, 66714117, 29331828, 35696052, 264509, 264605, 264906, 264907, 264606, 66712202, 264909, 265007, 265010, 264931, 264933, 5581038, 265011, 265016, 264700, 264682, 264784, 264853, 264369, 264706, 264766, 264789, 21906767, 21906768, 21906769, 264693, 18108374, 18108375, 264657, 264685, 264637, 264639, 264659, 22279000, 22279002, 264669
1814	88176047 (3627, 3628)	Novel Protein sim. GBank g13643508 (AC005385) - hypothetical protein [Arabidopsis thaliana]	UNCLASSIFIED		264488, 35696266, 22276988, 264092, 264094, 264259, 29331822, 29331824, 29331825, 60432289, 29331826, 29331827, 29331828, 284105, 264107, 52644045, 56182435, 265009, 60432289, 60433356, 87166474, 87168559, 264369, 264288, 21906765, 35695917, 265021, 265022, 33657023, 33657109, 18108374, 35696423, 264638, 35696266, 264482
1815	85286473 (3629, 3630)	Novel Protein sim. GBank g1177893gP2670CY44_RAT ADENYLATE CYCLASE, TYPE IV (ATP PYROPHOSPHATE-LYASE) (ADENYL CYCLASE)	slut		22276995, 264306, 264509, 264807, 264908, 264910, 265011, 264766, 264766, 264634, 264636
1816	83738845 (3631, 3632)	Novel Protein sim. GBank g1176623gP4-1846Y086_CAMEL - HYPOTHETICAL 93.9 KD PROTEIN T20B12.6 IN CHROMOSOME III	UNCLASSIFIED		18108394, 18108397, 264509, 264907, 264908, 264909, 265009, 264991, 265011, 265017, 264687, 264689, 265022, 264691, 18108362, 18108366, 18108370, 18108374, 18108379, 264635, 294557, 264564, 264567, 264488, 22278897, 22278899, 60433046, 29331822, 29331824, 60432289, 52644045, 6070831, 265017, 265018, 265019, 18108351, 264682, 52644229, 21906765, 21906767, 21906768, 52644150, 33657023, 33657109, 27488262, 18108370, 18108374, 80770354, 58182333, 22279002
1817	88056268 (3633, 3634)	Novel Protein sim. GBank g1387912femb(CA54370) - (AL031607) hypothetical protein [Schistosoma haematophy- somae]	Contains protein domain (PF00400) WD domain, G-beta repeat	histone	35696266, 60433356, 264758, 264369, 264686, 21906769, 264693, 264632
1818	88056775 (3635, 3636)	Novel Protein sim. GBank g1387912femb(CA54370) - (Z70310) predicted using GeneFinder. Similarity to Mouse anaphin (P/R Acc. No. S3771). cDNA EST EMBL 101923 comes from this gene; cDNA EST EMBL 032335 comes from this gene; cDNA EST EMBL 032723 comes from this gene; cDNA EST E..	Contains protein domain (PF00023) Ark repeat	transcription factor	
1819	87759572 (3637, 3638)	Novel Protein sim. GBank g10031865fembP_00571.1p1-HFP - Igoma HMG fusion partner	UNCLASSIFIED		22278995, 22278997, 22278998, 264478, 29331826, 264908, 265007, 265008, 265009, 60432289, 265017, 265018, 265019, 264448, 264288, 21906768, 21906769, 265020, 18108381, 18108384, 22279000, 22279002, 264567

1820	87769455 (3639, 3640)				264305, 264307, 264694
1821	80431510 (3641, 3642)				264705, 264706, 265197
1822	91221523 (3643, 3644)	Novel Protein sim. GBank gI4884130junc[CAB43272.1] - (AU050101) hypothetical protein [Homo sapiens]			2227897, 2227898, 2227899, 2227900, 2227901, 2227902, 2227903, 2227904, 2227905, 2227906, 2227907, 2227908, 264699, 26531824, 26531825, 26531826, 26531827, 26531828, 26531829, 26531830, 26531831, 264691, 264693, 26531830, 60170831, 264691, 264693, 60433356, 264696, 265017, 265019, 18108351, 264763, 264683, 21906765, 21906767, 21906768, 21906769, 35695917, 265020, 265021, 33657023, 18108364, 18108370, 35695855, 22279000, 22279002, 264486, 264759, 264311, 264288, 267768, 264693, 35696423, 264634, 18108365, 264486
1823	85522330 (3645, 3646)			UNCLASSIFIED	UNCLASSIFIED
1824	88612025 (3647, 3648)	Novel Protein sim. GBank gI47072iprI44018 - mucn 7 precursor, salivary - human	Contains protein domain (PF00047) - Immunoglobulin domain	UNCLASSIFIED	264907, 264608, 264909, 264511, 264631, 264634, 264635, 264637, 264638, 264639, 264758, 264760
1825	87430175 (3649, 3650)	Novel Protein sim. GBank gI075603junc[C441943] - (AU022372) hypothetical protein [Arabidopsis thaliana]		UNCLASSIFIED	60432949, 264910, 264487
1826	91723612 (3651, 3652)	Novel Protein sim. GBank gI668065iprIA027732 - 1(AF13295 - (AF132987) CGL-23 protein [Homo sapiens])		ATPase, associated	52644507, 52645105, 52646342, 22278984, 22278985, 35694075, 264293, 60432049, 52645090, 35696092, 60712502, 52644045, 265008, 265009, 60432228, 60433366, 60433438, 52646317, 52644296, 265011, 87168558, 264448, 264288, 264369, 264688, 52644229, 264689, 21906765, 21906768, 265020, 60170615, 52644150, 33657023, 27486262, 27486264, 27486265, 35695783, 35696423, 35695855, 83373004, 87168518, 264404, 22279002
1827	81647212 (3653, 3654)				264758

1123	95074017 (3655, 3656)	Novel Protein sim. GBank gH45033711eNP_001419.1 PENO1 - enolase 1, (alpha)	Contains protein domain (PF00113) Enolase	onco gene	264488, 52646842, 5612575, 22278986, 35595286, 22278987, 22278959, 264091, 264033, 60432045, 264459, 26331822, 26331824, 60741117, 26331823, 60432289, 26331826, 26331827, 26331828, 264106, 26418345, 26418346, 26418347, 60712302, 52644045, 50183433, 265009, 60170931, 6045412, 265007, 265008, 265009, 60170931, 6045412, 265007, 264533, 60433358, 60433438, 264768, 33109944, 21906764, 87168474, 265010, 265011, 87168555, 265017, 265019, 264761, 264762, 264444, 264764, 264583, 264288, 264369, 18108355, 264768, 18108357, 18108358, 264688, 264769, 264689, 21906768, 21906769, 35695917, 285021, 60170615, 33657023, 33657349, 263972, 55811576, 35695955, 264635, 264555, 264558, 264038, 264457, 87168518, 22279000, 22279002, 264563, 264482, 264555, 264484, 264457
1829	80197720 (3657, 3658)				264508, 264038, 264508, 264492, 29331827, 264509, 264510, 264511, 264512, 52645145, 22278984, 2718905, 55695686, 22278986, 22278987, 22278988, 22278989, 29331822, 29331825, 35695952, 52646317, 52644298, 87168559, 265019, 21906769, 21906768, 21906767, 21906768, 21906769, 35695917, 265021, 33657023, 52645120, 33657109, 33657182, 27486281, 27486282, 35695763, 263374, 35695423, 35695555, 52644332
1831	84138063 (3661, 3662)			UNCLASSIFIED	29331824, 35696052, 29331830, 264595, 264758, 265010, 265019, 265022, 264893, 85274791 264002
1830	94312942 (3659, 3660)	Novel Protein sim. GBank gH2246532 (U83072) - ORF 73, contains large complex repeat CR 73 [Kaposi's sarcoma- associated herpesvirus]		nuclease	52645145, 22278984, 2718905, 55695686, 22278986, 22278987, 22278988, 22278989, 29331822, 29331825, 35695952, 52646317, 52644298, 87168559, 265019, 21906769, 21906768, 21906767, 21906768, 21906769, 35695917, 265021, 33657023, 52645120, 33657109, 33657182, 27486281, 27486282, 35695763, 263374, 35695423, 35695555, 52644332
1832	84521863 (3663, 3664)	Novel Protein sim. GBank gH1330345 (U80750) - coded for by C. elegans cDNA yK3ab1.5, coded for by C. elegans cDNA yK13h10.5, coded for by C. elegans cDNA yK46a8.5, coded for by C. elegans cDNA yK46d5.5, coded for by C. elegans cDNA yK43c2.5, coded for by C. elegans cDNA yK46e8.....			

1833	g95314184 (3665, 3666)	Novel Protein sim. GBank g1517443(rlfnp_060626) protein kinase beta (CdkR-like)	Contains protein domain (PF00069) Eukaryotic protein kinase domain	kinase	22278994, 22278987, 22278988, 22278989, 264259, 29331822, 29331824, 60432289, 29331827, 35690692, 29148499, 264558, 264509, 264908, 264907, 68712502, 264908, 52844045, 264508, 264512, 265008, 264591, 264593, 60433358, 21906754, 33857094, 265011, 265017, 264604, 265016, 265019, 264594, 264595, 264596, 264597, 264655, 264656, 264657, 21906755, 21906756, 21905767, 21905768, 21906769, 20148939, 265020, 265021, 264690, 264892, 33857023, 65274620, 33857182, 27485284, 33857349, 65274791, 264634, 264635, 264556, 264557, 264558, 264559, 18108385, 56526486, 87168516, 60432113, 22279000, 22279002, 264563
1834	g0562730 (367, 368)				264259, 264607, 264885, 22279000, 22279002
1835	g4135718 (369, 3670)			UNCLASSIFIED	22278998, 29331822, 29331826, 87168474, 264603, 21906768, 263976, 35690695, 83373044
1836	g7348450 (3671, 3672)	Novel Protein sim. GBank g14759236(rlfnp_004268, 1pUCP4 - uncoupling protein 4	Contains protein domain (PF00153) Mitochondrial carrier protein	transport	29331822, 264908, 265019, 264794, 264686, 264687, 264688, 264689, 264690, 22278995, 29148499, 265026, 265028, 265009, 265010, 264683, 21906765, 29148827, 29148920, 265020, 265021, 265022, 65274620, 18108370, 18108374, 264556, 18108385
1837	g4324237 (3673, 3674)	Novel Protein sim. GBank g13334400(spD24574)(UPPE, DROME - UBIQUITIN CARBOXYL-TERMINAL HYDROLASE 64E (UBIQUITIN THIOLESTERASE 64E) (UBIQUITIN-SPECIFIC PROCESSING PROTEASE 64E) (DEUBIQUITINATING ENZYME 64E)	Contains protein domain (PF00443) Ubiquitin carboxyl terminal hydrolase family 2	ubiquitin	
1838	g4324269 (3675, 3676)	Novel Protein sim. GBank g13625599(jA56154 - Acl substrate ena (enabes) - fruit fly Drosophila melanogaster)	Contains protein domain (PF00568) - WH1 domain		29331822, 265017, 254760, 265020, 83373044
1839	g7456508 (3677, 3678)	Novel Protein sim. GBank g2117310(emjCAB09116.1) - (256020) hypocholesterol protein [Schizosaccharomyces	UNCLASSIFIED	UNCLASSIFIED	60433436, 294407, 21906765, 21906766, 265021, 33857109, 264556
1840	g7391708 (3679, 3680)	Novel Protein sim. GBank g11275601(spP2349MM10, MOUSE - PROTEIN MOV-10		UNCLASSIFIED	264683
1841	g5818445 (3681, 3682)	Novel Protein sim. GBank g14572464(hbA023834.1)(F12365 - (AF 12365) FEZ1 [Homo sapiens])			5616257, 29331824, 29331826, 60433356, 264686, 33857023, 263987, 18108370, 18108374, 204631, 204655, 264556, 264639

1852	93413170 (3703, 3704)	Novel Protein sim. GBank g15174629[eflPp_006900_lipMAS - protein inhibitor of activated STAT3]	UNCLASSIFIED	56102575, 35666206, 22278996, 22278997, 22278999, 264490, 60432049, 264259, 29331822, 29331824, 29331825, 29331826, 60432289, 29331827, 35696052, 52840405, 265007, 264510, 60432229, 60433356, 60433438, 5812038, 6527444, 265018, 265019, 10108351, 264448, 264686, 264687, 21007654, 21906767, 21906768, 265021, 21907653, 264686, 264687, 264688, 18108370, 18108371, 18108372, 18108373, 18108374, 18108375, 18108376, 35695855, 56102323, 60432113, 22279002, 264583, 264603, 264768, 52844507, 264769, 21908765, 21906767, 21906769, 22278995, 5694075, 22278999, 52844150, 264259, 264652, 29331822, 29331824, 52845129, 29331827, 35695970, 35657349, 35665763, 264508, 264906, 264628, 264607, 264629, 284909, 35696423, 35696555, 264510, 265006, 264511, 264512, 264630, 265009, 264631, 264910, 264634, 264635, 264637, 264593, 264638, 264639, 33657402, 18108385, 52846317, 52844296, 27168518, 37198559, 264602, 265017, 22279000, 265018, 264760, 264762, 264652, 264446, 264764, 264684, 264687, 264208, 264368, 264369, 264392
1853	91222267 (3705, 3706)	Novel Protein sim. GBank g1854065[mbi/CAAG3337] - (X33413) U88 [Human herpesvirus 6]	UNCLASSIFIED	
1854	86038152 (3707, 3708)	Novel Protein sim. GBank g12072964 (U33669) - relative p150 [Homo sapiens]	nuclease	
1855	91221459 (3709, 3710)	Novel Protein sim. GBank g14539520[mbi/AB3994.1] - (AL035424) dA22D12.1 (novel protein similar to Drosophila Ketch (Ring Canal protein, KCL) and a heterogeneous set of other types of proteins) [Homo sapiens]	Contains protein domain (PF01344) - nucd_rept Ketch motif	18105392, 52846365, 65274572, 56182575, 22278994, 56994075, 22278996, 35696256, 22278999, 264259, 60432049, 264908, 264510, 265007, 265008, 265009, 264595, 21908764, 87168474, 285011, 87166559, 264681, 264208, 264768, 18108359, 21007654, 21906768, 29148827, 285020, 265021, 265022, 52844150, 33657023, 33657109, 18108372, 18108374, 18108376, 35696423, 264631, 264696, 21008381, 264482
1856	94231871 (3711, 3712)	Novel Protein sim. GBank g1935477[mbi/CA06948] - (AL008279) acylglucosaminyltransferase-like protein [Mus musculus]	UNCLASSIFIED	36994075, 264259, 29331828, 264511, 264510, 264768, 264683, 264637, 18108381, 83373695, 264259, 264906, 80170831, 22278999
1857	94324455 (3713, 3714)	Novel Protein sim. GBank g142270[mbi/AD16120] - (AF094608) dentin phosphophoryn [Homo sapiens]	ATPase-associated	264448, 264686, 265020, 265022, 33657109, 60170394, 83373044

1658	87625311 (3715, 3716)	Novel Protein sim. GBank g14981903/gp1A036415.1(AE00178) - (AE00178) Ribosomal protein S15 [Thermotoga maritima]	Contains protein domain (PF00312) - Ribosomal protein S15	264757
1659	84407464 (3717, 3718)	Novel Protein sim. GBank g142403710/gp1A0437.11 - (A0502721) KIAA0914 protein [Homo sapiens]		22278996, 26331824, 265007, 23109064, 264591, 264638, 264639, 264640, 27486361, 52644332, 22279002 265019
1660	17625508 (3719, 3720)	Novel Protein sim. GBank g14039522 (AF099731) - coronin 3.1 [Homo sapiens]		
1661	80386370 (3721, 3722)	Novel Protein sim. GBank g02143037/gp1A04505 calcium- dependent actin-binding protein - rat	Contains protein domain (PF00285) Citrate synthase	264487, 264529, 26331822, 26331824, 29331825, 265007, 265009, 264591, 33109954, 265010, 265019, 264369, 264288, 264686, 264691, 264693, 27486284, 18108370, 18108374, 263977, 55911576, 59182323, 264639, 22279000, 22279002, 264482
1662	87372823 (3723, 3724)	Novel Protein sim. GBank g1254939p/p07313/KMCLC_RABIT - MYOSIN LG-IT CHAIN KINASE, SKELETAL MUSCLE (MLCK)	Contains protein domain (PF00069) - Eukaryotic protein kinase domain	35096286, 264299, 87164174, 264369, 21900766, 264336, 264363
1663	85775037 (3725, 3726)	Novel Protein sim. GBank g13326269/gp1A06999 - (A010842) Dof protein [Drosophila melanogaster]	UNCLASSIFIED	264601, 264768, 29148627, 29148629, 264769, 264770, 264771, 264772, 18108370, 264634, 264635, 264556, 264638, 18108387
1664	85474532 (3727, 3728)	Novel Protein sim. GBank g14222532/gp1A011966 - (AF071049) metalloproteinase CPX-1 [Mus musculus]	Contains protein domain (PF00754) F58 type C domain	22278998, 264480, 26331822, 66714117, 66712502, 265006, 265007, 265008, 265009, 264591, 04433433, 265010, 265019, 264760, 264448, 264768, 29148627, 29148629, 265020, 265022, 18108385, 60432113 18108374, 264769, 18108377, 1905765, 21900766, 35696423, 56182575, 21905769, 29148628, 35696286, 35699917, 285021, 29148629, 264511, 264512, 264534, 264535, 6070831, 52644150, 264555, 264691, 264259, 264556, 264692, 264557, 33857023, 60433500, 29331822, 264689, 264595, 29331824, 19108387, 18108388, 18108389, 18108390, 18108391, 33856970, 87164518, 265017, 60431602, 22279000, 264508, 264509, 18108351, 264987, 264682, 264567, 18108372, 284765, 264486 264094
1665	87740827 (3729, 3730)	Novel Protein sim. GBank g10496727/gp1A03673/Y265_HUMAN - HYPOTHETICAL PROTEIN KIAA0256		
1666	67268316 (3731, 3732)	Novel Protein sim. GBank g155262617/gp1A0345748.11 - (AL080157) hypothetical protein [Homo sapiens]	kinase	
1667	84579159 (3733, 3734)	Novel Protein sim. GBank g13559530 (AF078096) - forkheadwinged helix-like transcription factor 7 [Homo sapiens]	UNCLASSIFIED	

1878	55351056 (3755, 3765)	Novel Protein sim. GBank gi45103453pJAD21434.11 - (AC006921) unknown protein [Arabidopsis thaliana]	Contains protein domain (PF01428) - AN1-like Zinc finger	264569, 264488, 35656285, 56994075, 264259, 29331822, 29331824, 29331825, 35696052, 29331828, 29346456, 2646706, 2646707, 2646708, 2646709, 2646710, 2646711, 2646712, 265008, 264910, 33657402, 264934, 264758, 55912038, 265011, 2646802, 264768, 18108351, 264762, 264681, 264682, 264764, 264399, 264428, 264768, 264687, 264768, 264769, 21906766, 35695917, 265021, 60170815, 33657023, 264692, 264693, 33657109, 27486265, 264628, 18108370, 264629, 18108374, 35696423, 264634, 264635, 264555, 264638, 264639, 83373044, 18108395, 56526486, 87166516, 264953, 264554, 264565, 264486, 264567, 265005, 264907, 264908, 265007, 264565, 264586
1879	93310583 (3757, 3758)	Novel Protein sim. GBank gi4929643pJAD34082.11(AF15184 - (AF151445) CGI-87 protein [Homo sapiens])	UNCLASSIFIED	264766, 264691, 264692, 83373044
1880	91072978 (3759, 3760)	Novel Protein sim. GBank gi15507159pJCA4592831 - (Y08026) immune associated protein 38 [Mus musculus]	UNCLASSIFIED	264559, 264905, 264908, 264909, 264910, 264762, 264887, 33657023, 264632
1881	80214949 (3761, 3762)	Novel Protein sim. GBank gi931441pJH04055 - (U08026) immune associated protein 38 [Mus musculus]	UNCLASSIFIED	264559, 264905, 264908, 264909, 264910, 264762, 264887, 33657023, 264632
1882	85552450 (3763, 3764)	Novel Protein sim. GBank gi2384566 (AF023985) - No definition line found [Caenorhabditis elegans]	UNCLASSIFIED	264908, 21906766, 18108370, 263974, 87166518
1883	94216817 (3765, 3766)	Novel Protein sim. GBank gi13512181pJF72281TES2_MOUSE - TESTIN 2 (TES2) [CONTAINS: TESTIN 1 (TES1)]	Contains protein domain (PF00412) - LIM domain containing proteins	264908, 264910, 87185559, 21906766, 264636

1891	87013905 (3791, 3792)			UNCLASSIFIED	264696, 264788, 264687, 264692, 264693, 26331822, 26331824, 264508, 264405, 264636, 18108370, 264628, 264507, 264908, 264699, 18108379, 265007, 265008, 264510, 264627, 18108391, 264603, 264596, 18108384, 265010, 265011, 264601, 264602, 264563, 264569
1892	37642825 (3793, 3794)	Novel Protein sim. GBank g15699535d0JBAA3005111- (A90290022) KIAA1099 protein (Homo sapiens)	Contains protein domain (PF01412) - putative GTP-ase activating protein for Arf	UNCLASSIFIED	22278895, 264599, 27168559, 18108351, 284448, 264982, 264020, 264683, 18108374, 22279000
1893	86533826 (3795, 3796)		lyminin	UNCLASSIFIED	264569, 60274572, 22278697, 22278699, 264259, 29331822, 29331824, 66714117, 29331826, 264906, 265006, 265008, 265009, 264592, 265018, 264681, 264448, 264683, 18108354, 264589, 264684, 264685, 264766, 264687, 264689, 21906768, 265020, 265022, 60170615, 26544150, 264690, 264691, 264692, 33657023, 264693, 33657109, 264628, 18108374, 35695985, 264630, 264632, 264834, 264557, 264558, 60170394, 18108381, 18108385, 22279000
1894	86989120 (3797, 3798)			UNCLASSIFIED	264508, 264905, 264906, 264607, 264564, 56182375, 265009, 264653, 264655, 56182376, 265010, 264653, 264655
1895	87831801 (3799, 3790)	Novel Protein sim. GBank g15262542amb1CAB45726111- (A1801133) hypothetical protein [Homo sapiens]	Contains protein domain (PF00435) - Spectrin repeat	UNCLASSIFIED	264107, 264905, 264908, 264910, 2650031, 264758, 265010, 265018, 264448, 264680, 264768, 33657109, 264628, 55810764, 18108379, 264634, 56182333, 56526486
1896	85673555 (3791, 3792)		UNCLASSIFIED	UNCLASSIFIED	264607, 265008, 264682, 264686, 21906768, 264629, 264631, 264634, 284555, 264729
1897	80656569 (3793, 3794)	Novel Protein sim. GBank g1288305p3919JALUG_HUMAN - III ALU SUBFAMILY SP - WARNING ENTRY IIII	cadherin	UNCLASSIFIED	22278696, 22278698, 22278699, 26331824, 26331825, 60432288, 29331827, 35696092, 29331828, 265008, 265010, 264681, 264682, 264448, 264369, 52644224, 21906765, 21906766, 21906768, 21906769, 60170615, 55810764, 22279000
1898	87617637 (3795, 3796)	Novel Protein sim. GBank g1275605p39224JMV10_MOUSE - PROTEIN MOV-10	helicase	UNCLASSIFIED	264599, 264508, 264509, 60432228, 264765, 21906765, 21906766, 21906768, 21906769
1899	86672007 (3797, 3798)	Novel Protein sim. GBank g12908519 (A1031548) - erythrocyte membrane glycoprotein [Homo sapiens]	glycoprotein	UNCLASSIFIED	264599, 264508, 264509, 60432228, 264765, 21906765, 21906766, 21906768, 21906769
1900	87641853 (3799, 3800)	Novel Protein sim. GBank g14102881 (A1017250) - vitellinogen precursor [Odocoileus arcticus]	Ammonium Transporter Family	UNCLASSIFIED	264599, 264508, 264509, 60432228, 264765, 21906765, 21906766, 21906768, 21906769

1901	95196647 (3801, 3802)	Novel Protein sim. GBank g 545589 g 367 B561A_EAT SUBUNIT PROTEIN TRANSPORT PROTEIN SEC1 ALPHA	Contains protein domain (PF00242) DNA polymerase (Nta) N-terminal domain	transport	264483, 32644307, 506456156, 18108356, 52546365, 52666542, 18103397, 36182515, 2278784, 22278855, 36940703, 3569036, 2278784, 22278855, 36940703, 3569036, 264259, 39331022, 52645090, 26531824, 26531823, 60714117, 26531826, 60432288, 26531827, 26531828, 36960052, 33656970, 29146498, 264906, 264907, 26531830, 264906, 52644045, 264909, 264112, 265006, 264512, 265006, 264910, 265009, 60170831, 60432229, 60433356, 33657402, 60433438, 55812038, 264758, 33109954, 21906754, 33657084, 52642396, 87168474, 265010, 265011, 87168555, 265017, 265018, 265019, 18100351, 264448, 264288, 264686, 52644223, 21906765, 21906766, 21906767, 21906769, 55811957, 36969917, 265020, 265021, 32644150, 18103622, 33657023, 264683, 363967, 33657109, 33657162, 264684, 33657149, 33657153, 18100370, 18100370, 33657149, 33657153, 33657155, 60431850, 264638, 263981, 52643332, 60170394, 83372044, 18108365, 87168518, 60432113, 264564, 264107, 2653976
1902	80202013 (3803, 3804)	Novel Protein sim. GBank g 4426613 g A4D20451 - (AF088786) SLM-1 [Mus musculus]		dna_jna_bind	264259, 26531825, 26531827, 264508, 264907, 265006, 60170831, 60433356, 60433438, 264759, 21906754, 264448, 264288, 265021, 265022, 33657023, 264683, 55811576, 264555, 264585, 22779000
1903	87778554 (3805, 3806)	Novel Protein sim. GBank g 3747107 (AF057541) - unknown [Rattus norvegicus]		UNCLASSIFIED	264259, 26531825, 26531827, 264508, 264907, 265006, 60170831, 60433356, 60433438, 264759, 21906754, 264448, 264288, 265021, 265022, 33657023, 264683, 55811576, 264555, 264585, 22779000
1904	80434213 (3807, 3808)	Novel Protein sim. GBank g 1332871 g P47147Y J80_YEAST - HYPOTHETICAL 812 KD PROTEIN IN CPAP-NF1 INTERGENIC REGION (AB011167) KIAA0595 protein [Homo sapiens]		groud	264509, 264905, 264906, 264907, 264908, 265007, 264910, 264906, 264768, 264687, 264769, 264693, 264628, 18103374, 264634, 264636, 264637, 264635, 264638, 33657023, 26531824, 33657109, 18103365, 33657023, 26531824, 33657109, 29146498, 264906, 264909, 264605, 264906, 264907, 60172502, 264908, 264909, 35898423, 35895555, 264510, 264511, 264512, 264910, 264634, 264635, 264637, 264638, 33657402, 264758, 85658542, 264602, 264760, 264761, 264482, 264563, 264762, 264483, 264764, 264586, 264288, 264756
1905	95351140 (3809, 3810)	Novel Protein sim. GBank g 304374 g B433521 - (AB011167) KIAA0595 protein [Homo sapiens]	Contains protein domain (PF00243) - Bacterial muT protein		264509, 264905, 264906, 264907, 264908, 265007, 264910, 264906, 264768, 264687, 264769, 264693, 264628, 18103374, 264634, 264636, 264637, 264635, 264638, 33657023, 26531824, 33657109, 18103365, 33657023, 26531824, 33657109, 29146498, 264906, 264909, 264605, 264906, 264907, 60172502, 264908, 264909, 35898423, 35895555, 264510, 264511, 264512, 264910, 264634, 264635, 264637, 264638, 33657402, 264758, 85658542, 264602, 264760, 264761, 264482, 264563, 264762, 264483, 264764, 264586, 264288, 264756
1906	12763822 (3811, 3812)			UNCLASSIFIED	264259, 26531825, 26531827, 264508, 264907, 265006, 60170831, 60433356, 60433438, 264759, 21906754, 264448, 264288, 265021, 265022, 33657023, 264683, 55811576, 264555, 264585, 22779000

1807	5535114 (3813, 3814)	Novel Protein sim. GBank g1492585gipAC04033 (AF15181 - (AF151816) CGI-58 protein [Homo sapiens])	Contains protein domain (PF00561) - aliphatic hydrolase fold	UNCLASSIFIED	65274572, 22278996, 16590686, 22278998, 22278999, 264259, 60432048, 29331822, 29331823, 60432049, 29331828, 35695002, 264593, 60433356, 21906754, 55811386, 85655542, 87165559, 265016, 264581, 264682, 264684, 264288, 21906765, 21906766, 21906768, 265020, 265022, 264690, 52644150, 264692, 33657023, 264693, 33657109, 35695955, 264636, 264638, 60432113
1808	95313841 (3815, 3816)	Novel Protein sim. GBank g13936770 (AF109906) - NG22 [Mus musculus]		UNCLASSIFIED	264468, 65274572, 5618257, 56181686, 22278996, 22278997, 22278998, 22278999, 60432049, 264259, 29331822, 60432038, 29331828, 29331827, 29331823, 29146498, 264685, 264696, 264697, 264698, 29146499, 264699, 264700, 264701, 264702, 264703, 264704, 264705, 264706, 264707, 264708, 264709, 264710, 60170831, 264592, 60433366, 33657402, 264594, 60433438, 264596, 55812038, 33109954, 52646317, 265011, 265017, 264604, 265018, 265019, 264605, 55811150, 264681, 264446, 264288, 264686, 264688, 264769, 21906765, 21906766, 21906767, 21906768, 21906769, 55811957, 26148629, 35695917, 265020, 265022, 264691, 264692, 18103354, 65274620, 33657109, 33657345, 35695763, 18108374, 263978, 55810764, 55811576, 35696423, 65274791, 264631, 264632, 264556, 264557, 60170394, 56182323, 53373044, 18108385, 60432113, 22278900, 22278902, 264596, 264597, 264598, 264599, 264600, 264601, 264602, 264603, 264604, 264605, 264606, 264607, 264608, 264609, 264610, 264611, 264612, 264758, 264764, 264766, 18108370, 264634, 264637, 264486
1909	85514505 (3817, 3818)	Novel Protein sim. GBank g12224653bj[BL020813] - (A8002354) KIAA0036 [Homo sapiens]		UNCLASSIFIED	35695286, 22278996, 22278998, 35698052, 264509, 264905, 264906, 264907, 264908, 264909, 264910, 265007, 264512, 264910, 264909, 264511, 265007, 264512, 264910, 264758, 265011, 264601, 264602, 264604, 264605, 264761, 264764, 264288, 264768, 264769, 264687, 264769, 35695917, 265021, 52644150, 264692, 264628, 18108370, 264629, 18108372, 18108374, 35695855, 264631, 264634, 264635, 264636, 264637, 264638, 18108385, 264595, 264596, 264486
1910	54216521 (3819, 3820)	Novel Protein sim. GBank g11581218upg728TIES2 MOUSE - TESTIN2 (TIES2) [CONTAINS: TESTIN 1 (TES1)]	Contains protein domain (PF00412) - LIM domain containing proteins	homeobox	

1911	91725345 (3827, 3828)	Novel Protein sim. GBank gl4808359 glA030784.1 AC06653 - (AC06653) hypomethyl protein [Homo sapiens]	Contains protein domain (PF01119) - DNA mismatch repair protein	nuclease	18108394, 56182575, 56182181, 29331826, 29331827, 33659970, 264906, 265007, 264591, 55812035, 87168593, 264446, 264592, 264593, 18108395, 55818282, 264899, 264903, 18108395, 55818156, 264556, 18108395, 18108399
1912	95413519 (3823, 3824)	Novel Protein sim. GBank gl568439 db JBA43003.1 - (AB028974) KIAA1051 protein [Homo sapiens]	Contains protein domain (PF00039) - Zinc finger, CCHC class	UNCLASSIFIED	18108397, 56182575, 56181686, 22278994, 22278995, 56994075, 22278996, 22278998, 284259, 29331822, 29331824, 56182181, 29331825, 6671417, 35696052, 264905, 264906, 264907, 264908, 55844045, 56182435, 265007, 265008, 264910, 265009, 264591, 264596, 85274444, 55811388, 87168474, 265011, 87168599, 265018, 265019, 264760, 18108351, 264681, 264369, 264684, 264288, 264686, 264768, 2190765, 2190766, 2190767, 2190768, 2190769, 59811957, 265020, 265021, 265022, 60170515, 264692, 33657023, 264693, 18108376, 55811576, 35696423, 63274791, 55811577, 35696424, 35696425, 35696426, 22792002, 264683, 2194588
1913	95309546 (3825, 3826)	Novel Protein sim. GBank gl5032245 ref NP_005655.1 p2NF2 - zinc finger protein (C2H2) homologous to mouse MOK-2			56182575, 22278994, 22278995, 56994075, 22278996, 22278998, 22278999, 29331828, 29331827, 265008, 55812038, 265010, 265017, 265019, 265019, 264681, 18108351, 264683, 264764, 264599, 264288, 264685, 264686, 264769, 2190765, 2190766, 2190768, 2190769, 55811957, 265020, 265022, 264691, 55811576, 264694, 264635, 264639, 56182323, 83373044, 18108395
1914	83423982 (3827, 3828)	Novel Protein sim. GBank gl4398604 db JBA47682.1 - (AB023197) KIAA0980 protein [Homo sapiens]	Contains protein domain (PF00036) - EF hand	struct	56182575, 29331824, 35696052, 264906, 264908, 264928, 264909, 264592, 264798, 87168599, 18108351, 18108354, 264684, 264686, 33657023, 264693, 264694, 264695, 264696, 264697, 264698
1915	95304459 (3829, 3830)	Novel Protein sim. GBank gl568431 gl JBA432981.1 - (AB028982) KIAA1039 protein [Homo sapiens]		UNCLASSIFIED	264259, 29331824, 29331826, 29331827, 264593, 264609, 265009, 265017, 265019, 264768, 264769, 264699, 264628, 264635, 264637, 264639, 83373044, 264595
1916	79640761 (3831, 3832)				264683, 264639

1027	87641863 (3443, 3844)	Novel Protein sim. GBank gi138595pp024545V172.CHICK - VITELLOGENIN II REPLICATOR (MAJOR VITELLOGENIN) [CONTAINS LPOVITELIN I (LVI); PHOSVITIN (PV); LPOVITELIN II (LVL); YCP40]		UNCLASSIFIED	264686, 264688, 264690, 1810370, 264909, 1810374, 265008, 264637, 264584, 1810331
1028	94323559 (3345, 3846)	Novel Protein sim. GBank gi1191101p020211EBN1_EBV - EBNA1 NUCLEAR PROTEIN		UNCLASSIFIED	264438, 264489, 22278956, 264094, 264259, 35698052, 264509, 264905, 264908, 264907, 264908, 68712502, 264909, 264511, 264512, 265007, 264910, 265008, 264591, 264592, 264593, 264594, 264595, 264758, 264596, 264759, 265011, 265017, 265018, 265019, 5581150, 264081, 264162, 264448, 264794, 264288, 264369, 264766, 264767, 264686, 264687, 264768, 264769, 56181562, 264689, 21906766, 264681, 33657023, 264693, 65274620, 33657109, 18103370, 264628, 264629, 35699423, 264630, 264631, 264632, 264634, 264555, 264638, 294637, 264638, 264839, 264558, 33373044, 87188518, 264593, 264594, 264595, 264596, 264597, 26331826, 264906, 264908, 264939, 265017, 265018, 265019, 265020, 264681, 264682, 264637, 18100385, 264605
1029	87330925 (3047, 3848)	Novel Protein sim. GBank gi1387855jms1CA966571 - (Z72511) possible zinc finger protein: cDNA EST EMBL.M89115 comes from this gene: cDNA EST EMBL.D71533 comes from this gene: cDNA EST EMBL.D72314 comes from this gene: cDNA EST EMBL.D75164 comes from this gene: cDNA EST EMBL.C1...	Contains protein domain (PF00097) - Zinc finger, C2H2 type (RING finger)		
1030	87628338 (3849, 3850)	Novel Protein sim. GBank gi14981003giAACD36415.1(AE00178) - ribosomal protein S15 [Thermotoga maritima]	Contains protein domain (PF00312) - ribosomal protein S15		
1031	88094739 (3851, 3852)	Novel Protein sim. GBank gi12246532 (U93872) - ORF 73, contains large complex repeat CR 73 [Kapoor's sarcoma- associated herpesvirus]			
1032	85654657 (3853, 3854)	Novel Protein sim. GBank gi13034332 (giJBA26-80) - (AB011128) KIAA0554 protein [Homo sapiens]	stuct		
1033	87790054 (3855, 3856)	Novel Protein sim. GBank gi11855765 (giJBA13377) - (D87433) KIAA0248 [Homo sapiens]			
1034	86997236 (3857, 3858)	Novel Protein sim. GBank gi15001993giAAC037247.1(JAF13432) - (AF134321) c1emic AFGPhylogenetic-like serine protease precursor [Dissostichus nasutus]	Contains protein domain (PF00193) - Extracellular link domain	UNCLASSIFIED	264508, 264591, 33657102, 265017, 264768, 264632, 264658, 264639

1930	87859128 (3859, 3860)	Novel Protein sim. GBank g1170220qplP25963jNBL4_MOUSE - NBL 4 PROTEIN		phosphatase	35895286, 29331823, 254905, 264907, 264908, 264909, 264511, 264910, 264758, 264601, 265017, 265019, 264605, 264760, 264764, 264765, 264688, 264789, 265022, 35896427, 264639, 26432113
1931	87782729 (3861, 3862)	Novel Protein sim. GBank g1404634 (J01840) - Sema4 intracellular kinase [Mus musculus]		kinase	264906, 264908, 80432229, 264758, 264764, 264765, 265020, 264622, 264634, 264637, 264639, 264651, 264655
1932	15030972 (3863, 3864)			UNCLASSIFIED	
1933	11613868 (3865, 3866)	Novel Protein sim. GBank g115748dbjBAA36494 - (AB022023) nonmuscle myosin heavy chain B [Bos taurus]		strand	58182575, 58182435, 264610, 264757, 264758, 55812038, 55811386, 265018, 55811150, 21906765, 264681, 264631, 264635, 264637
1934	84438350 (3867, 3868)			UNCLASSIFIED	264686, 265011, 264511, 264905, 18108351, 264554, 264681, 264429, 18108370, 264566, 264764, 264365, 264495
1935	87782511 (3869, 3870)				80432289, 265057, 265010, 265011, 265018, 33657109, 18109374
1936	95414338 (3871, 3872)	Novel Protein sim. GBank g1482704qrefJNP_005110.1p1TRAP - thyroid hormone receptor-associated protein, 150 kDa subunit		kinase	85858542, 21906767, 35692977, 80170015, 264653, 33657109
1937	94847141 (3873, 3874)	Novel Protein sim. GBank g153187pJ53771 - ankyrin, erythrocyte - mouse		collagen	264488, 28146498, 264905, 264559
1938	87403277 (3875, 3876)	Novel Protein sim. GBank g1P64513pJ4022340.1AC00695 - (AC006956) hemoglobin, human		UNCLASSIFIED	86374572, 22278899, 60433289, 56182435, 80433345, 265017, 21906765, 21906766, 21906768, 55811957, 27488264, 35686423, 60432113, 264684
1939	81004978 (3877, 3878)	Novel Protein sim. GBank g1509558bjBAA03210 - (D14188) 50kDa lectin [Bombix mori]			264488, 29331822, 264448, 264683, 264286, 265020, 33657023, 264631
1940	87348510 (3879, 3880)	Novel Protein sim. GBank g11946300emb(CAA73132) - (Y12529) hypothetical protein [Silene latifolia]		UNCLASSIFIED	56594075, 22278898, 264429, 26531824, 29331827, 264905, 265008, 33657084, 265017, 265016, 264286, 264887, 21906765, 21906766, 21906767, 265020, 52644150, 27488264, 83373044, 18108387, 60432113, 22279002, 264565
1941	94147177 (3881, 3882)	Novel Protein sim. GBank g14206388 (AF005070) - rig-1 protein [Mus musculus]			264488, 18108398, 29331825, 27488267, 264509, 18108370, 18103374, 264682
1942	87641870 (3883, 3884)	Novel Protein sim. GBank g1492720qplA033048.1NC13391 - (AF133911) ARL-6 regulating protein [Mus musculus]		kinase	22781898, 29331822, 29331827, 35656492, 264511, 265009, 264492, 80432229, 265017, 265018, 265019, 264684, 264682, 33657109, 85274791, 264636
1943	94322288 (3885, 3886)	Novel Protein sim. GBank g1012882qpl015786jTIPD_DICD1 - TIPD PROTEIN			

1544	84232658 (3887, 3888)	Novel Protein sim. GBank g11708370e0j0AA13432) - (C87871) TIP120 [Rattus norvegicus]	UNCLASSIFIED	65374572, 22778994, 22778995, 22778996, 22778998, 22778999, 264259, 52945080, 28331824, 29331826, 29331827, 66712502, 56184435, 264512, 264910, 80170831, 26452293, 60433356, 33657402, 60433438, 26452293, 60433356, 33657402, 60433438, 87168558, 2650174, 87168474, 87168558, 2650174, 26452293, 60433356, 33657402, 60433438, 18108351, 264369, 264486, 264768, 21905765, 21905766, 21905767, 21905769, 35695917, 265020, 265021, 60170615, 264692, 33657023, 18108370, 18108374, 35696423, 35695555, 264634, 60170394, 264639, 83373044, 18108385, 55526486, 87168518, 60432113, 22279000, 264553, 264488, 22278996, 264510, 264511, 18108351, 264693, 264486, 264567
1545	87641872 (3889, 3890)	Novel Protein sim. GBank g14527204g0AA033048 (AF13391) ARL-6 [Rattus norvegicus]	UNCLASSIFIED	60432289, 25331827, 35696052, 265007, 285008, 60433356, 60434438, 264389, 56181562, 21905767, 52844150, 264693, 2532694, 264837, 87108518, 264583, 264512, 264528, 264768, 264698, 21906768, 264764, 264768, 264698, 21906768, 264693, 18108374, 35695955, 264635, 264693, 18108374, 35695955, 264635, 264837, 264638, 18108385, 22279002, 264909, 60170831, 264561, 264594, 235070, 265011, 264764, 264389, 254689, 264631, 264638
1546	87443950 (3891, 3892)	Novel Protein sim. GBank g124981045p027589AD20 BOVIN-ADRENAL MEDULLA 50 KD PROTEIN	UNCLASSIFIED	264389
1547	56438862 (3893, 3894)	Novel Protein sim. GBank g13914801lsp054889lP42 RAT - DNA-DIRECTED RNA POLYMERASE 1 135 KD POLYPEPTIDE (RNA POLYMERASE 1 SUBUNIT 2) (RPA135) (RNA POLYMERASE 1 127 KD SUBUNIT)	UNCLASSIFIED	264389
1548	95199174 (3895, 3896)	Novel Protein sim. GBank g15420387emb(CA946673.1) - (AJ243459) proteophosphoglycan [Leishmania major]	UNCLASSIFIED	264389
1549	7640128 (3897, 3898)	Novel Protein sim. GBank g15076766emb(CA93468.1) - (Z89627) predicted using Cestander, Similarity to E.coli proteolipid protein T4C (SWTCA_C_ECOLI) [Candida albicans]	UNCLASSIFIED	264389
1550	87788531 (3899, 3900)	Novel Protein sim. GBank g12626753dbj(BA234241) - (AB008742) sulfate transporter [Arabidopsis thaliana]	UNCLASSIFIED	264389
1551	96588253 (3901, 3902)	Novel Protein sim. GBank g14527204g0AA033048 (AF13391) ARL-6 [Rattus norvegicus]	UNCLASSIFIED	264389
1552	87069775 (3903, 3904)	Novel Protein sim. GBank g14829633g0AA034077.1 (AF15184) - (AF15184) COH-2 protein [Homo sapiens]	UNCLASSIFIED	264389

1853	20470371 (3805, 3806)	Novel Protein sim. GBank g11188715spj31721C1QB_RAY - COMPLEMENT C1q SUBCOMPONENT, B CHAIN PRECURSOR	Contains protein domain (PF00368) C1q domain	complement	264250, 264458
1854	91226025 (3807, 3808)	Novel Protein sim. GBank g14202710ds1BAA74914.1) - (AB020690) KIAA0891 protein [Homo sapiens]	Contains protein domain (PF00443) Ubiquitin carboxyl-terminal hydrolase family 2	ubiquitin	264488, 265094, 18103394, 35696286, 22278958, 29331822, 66714117, 29331826, 29331827, 264508, 264509, 264505, 264506, 264907, 264908, 264909, 264510, 265006, 265007, 265008, 264910, 265009, 60176831, 90432229, 265011, 265017, 264603, 264604, 265019, 18108351, 264635, 264766, 264769, 35959917, 265020, 265021, 264691, 33557023, 264692, 33657109, 264628, 18103374, 35696423, 55811576, 35696555, 264630, 264635, 264636, 264638, 264639, 264640, 264641, 264642, 264643, 264644, 264645, 264646, 264647, 264648, 264649, 63372358, 264650, 264651, 264652, 264653, 264654, 264655, 264656, 264657, 264658, 58182575, 58182576, 22278904, 22278909, 264259, 29331822, 58182575, 29331824, 29331825, 29331826, 29331827, 35696052, 264508, 29331830, 265008, 265009, 264591, 55812038, 87168474, 265017, 265018, 265019, 264448, 264766, 21906765, 21906766, 21906767, 55811957, 265020, 265021, 52645128, 33657109, 27486264, 33857343, 35695763, 60431528, 18108374, 55811576, 35695855, 264635, 60431850, 264639, 83373044
1855	95308310 (3809, 3910)		UNCLASSIFIED		
1856	95032121 (3911, 3912)	Novel Protein sim. GBank g11665821db1BAA13407) - (D87469) Similar to D melanogaster cadherin-related tumor suppressor [Homo sapiens]	Contains protein domain (PF00028) Cadherin domain	cadherin	264488, 58182575, 22278905, 22278909, 264630, 29331822, 29331824, 26331825, 29331826, 29331827, 35696052, 264508, 264906, 264908, 264909, 264510, 265006, 264910, 264509, 55812038, 265010, 265018, 18103351, 264768, 55811562, 264639, 21906768, 21906769, 265022, 264691, 264628, 18108374, 55810764, 55811576, 35695855, 264631, 264632, 264635, 264637, 264639, 60170394, 58182323, 83373044, 18108395, 22278900, 22279002, 264653, 264554

1957	94326510 (3913, 3914)	Novel Protein sim. GBank gi 53867462 g BA276565.1 - (AB023229) KIA1012 protein [Homo sapiens]	UNCLASSIFIED	52646842, 56182575, 22278997, 22276998, 22278999, 29331824, 66741117, 29331827, 23146498, 264593, 33657402, 33109954, 97108474, 265018, 264445, 264389, 264288, 21906769, 21906766, 21906767, 21906768, 21906769, 21906769, 21906769, 21906769, 65274630, 27446264, 264521, 264592, 33659585, 22779003, 264482
1958	95513502 (3915, 3916)	Novel Protein sim. GBank gi 24022789 g BA276922.1 - (AB020676) KIA00869 protein [Homo sapiens]	UNCLASSIFIED	66432049, 29331824, 56182181, 66741117, 264107, 264109, 264909, 264511, 60170831, 60432229, 21906754, 265010, 21906769, 33659597, 265022, 65274620, 263367, 263976, 35966423, 264631, 264632, 264634, 264635, 18108385, 22779002, 22779002, 264593, 265019
1959	85701470 (3917, 3918)	Novel Protein sim. GBank gi 2281083 g CAB10860 - (Z98056) hypothetical protein [Schistosoma mansoni]	ubiquitin	264905, 264806, 264907, 264908, 264909, 265006, 265007, 264910, 264955, 265017, 264904, 265018, 18108391, 264164, 264389, 264769, 264635, 264765, 18108389, 264629, 18108379, 264635, 264635, 264637, 264638, 264486
1960	80308608 (3919, 3920)	Novel Protein sim. GBank gi 2274851 g BA215119 - (D64159) 3-7 gene product [Homo sapiens]	ubiquitin	264635
1961	16292657 (3921, 3922)	Novel Protein sim. GBank gi J721653 g BA33381 - (AB072933) acyl-CoA synthetase 5 [Rattus norvegicus]	UNCLASSIFIED	65274572, 264592, 264593, 265019, 264631, 65274572, 18108398, 35966266, 26331825, 60432289, 29331827, 264828, 265006, 265009, 60433356, 60433438, 21906754, 265020, 265021, 33657023, 33657109, 27486265, 35959585, 264555
1962	94317605 (3927, 3928)	Novel Protein sim. GBank gi 5562638 g CAB14752.1 - (AL080169) hypothetical protein [Homo sapiens]	cadherin	264488, 264902, 264255, 264400, 264905, 264906, 264907, 264908, 264909, 264510, 264511, 265007, 265008, 264910, 264992, 264593, 264594, 264595, 264796, 264600, 264603, 264604, 264605, 264760, 264762, 264764, 264765, 264766, 264767, 264768, 264769, 264770, 21906766, 264591, 264592, 264633, 18108370, 264593, 264594, 18108374, 264630, 264631, 264632, 264636, 264637, 264638, 18108382, 83373044, 18108385, 264483, 264634, 264665, 264666, 264486, 264567
1965	94317445 (3929, 3930)	Novel Protein sim. GBank gi 470171 g BA35294 - (AB01773) PEN-6 [Drosophila melanogaster]	ubiquitin	264488, 264510, 264760, 264768, 264486

1975	94316479 (3949, 3950)			UNCLASSIFIED	264488, 6071417, 26331836, 26331828, 55182435, 265006, 264357, 55812038, 265010, 265017, 264369, 55819373, 65274791, 34695655, 55182323, 60432113, 264255, 36596052, 265018, 265020, 265021, 33657109, 56526486
1976	95358914 (3951, 3952)			UNCLASSIFIED	264908, 264596, 365021, 264566
1977	94852864 (3953, 3954)	Novel Protein sim. GBank g 2495286sp 007702 NASU_RAT - SODIUM/SULFATE COTRANSPORTER (NA+/SULFATE COTRANSPORTER)	homeobox	transcription factor	60170831, 264566
1978	87441645 (3955, 3956)	Novel Protein sim. GBank g 10342 prip A3471 - Novel transcription factor KTF1 - flut ty (Orasipalla)		transcription factor	26331826, 29146498, 264905, 264907, 265003, 265010, 265016, 264886, 265017, 265018, 264904, 36595917, 265020, 60170915, 264693, 18108370, 264631, 264635, 264555, 264558, 18108394, 264631, 264635, 18108395, 87168516
1979	87627709 (3957, 3958)	Novel Protein sim. GBank g 224483 gpcu C485 (0238, 11 - meta-opsin) (flut ty (Orasipalla)) (297336) hypothetical protein (Arabidopsis thaliana)		UNCLASSIFIED	22278995, 22278997, 22278999, 264259, 29331825, 29331828, 29146498, 29146499, 264107, 264908, 264910, 264595, 21906754, 265010, 265017, 265018, 265019, 264448, 264288, 21906767, 3657023, 27486264, 18108370, 18108374, 18108376, 264630, 264631, 264635, 18108395, 87168516
1980	86577059 (3959, 3960)	Novel Protein sim. GBank g 4759290 ref NP_094642.1 p USP-1 - Ubiquitin carboxyl-terminal hydrolase, X-linked (AF-131149) Unknown [Homo sapiens]	ubiquitin	UNCLASSIFIED	22278995, 22278997, 22278999, 264259, 29331825, 29331828, 29146498, 29146499, 264107, 264908, 264910, 264595, 21906754, 265010, 265017, 265018, 265019, 264448, 264288, 21906767, 3657023, 27486264, 18108370, 18108374, 18108376, 264630, 264631, 264635, 18108395, 87168516
1981	87606974 (3961, 3962)	Novel Protein sim. GBank g 4759290 ref NP_094642.1 p USP-1 - Ubiquitin carboxyl-terminal hydrolase, X-linked (AF-131149) Unknown [Homo sapiens]		UNCLASSIFIED	22278995, 22278997, 22278999, 264259, 29331825, 29331828, 29146498, 29146499, 264107, 264908, 264910, 264595, 21906754, 265010, 265017, 265018, 265019, 264448, 264288, 21906767, 3657023, 27486264, 18108370, 18108374, 18108376, 264630, 264631, 264635, 18108395, 87168516
1982	90595267 (3963, 3964)	Novel Protein sim. GBank g 5688523 gpcu J3A433045.11 - (AB029016) KIAA1093 protein [Homo sapiens]		UNCLASSIFIED	22278995, 22278997, 22278999, 264259, 29331825, 29331828, 29146498, 29146499, 264107, 264908, 264910, 264595, 21906754, 265010, 265017, 265018, 265019, 264448, 264288, 21906767, 3657023, 27486264, 18108370, 18108374, 18108376, 264630, 264631, 264635, 18108395, 87168516
1983	95098668 (3965, 3966)	Novel Protein sim. GBank g 3417297 AC002310 - Unknown gene product [Homo sapiens]	Contains protein domain (PF 00096) Zinc finger, C2H2 type	transcription factor	22278995, 22278997, 22278999, 264259, 29331825, 29331828, 29146498, 29146499, 264107, 264908, 264910, 264595, 21906754, 265010, 265017, 265018, 265019, 264448, 264288, 21906767, 3657023, 27486264, 18108370, 18108374, 18108376, 264630, 264631, 264635, 18108395, 87168516

1984	85760099 (3967, 3968)	Novel Protein sim. GBank gi22969595(jb CAA11714.1) - (AL021887) hadD14 [Mycobacterium tuberculosis]	synthase	264688, 21900706, 5811927, 56994075, 265020, 255021, 22278999, 265022, 264259, 29331822, 39537182, 39740499, 264628, 18108370, 264908, 264623, 5811576, 3959553, 265006, 265007, 2649911, 265009, 26497094, 265010, 265017, 265019, 264238, 261078, 58115176, 264637, 58182323, 18108385, 264584
1985	85636887 (3969, 3970)	Novel Protein sim. GBank gi571213(jb AA047375.1) [AF12049] DEM1 protein [Homo sapiens]	glycoprotein	264760, 264238, 261078, 58115176, 264637, 58182323, 18108385, 264584
1986	80200507 (3971, 3972)	Novel Protein sim. GBank gi488443(jb AA031319.1) [AF14457] - (AF144573) Mx-interacting protein kinase PKM [Mesocricetus auratus]	UNCLASSIFIED	264488, 264629
1987	87011117 (3973, 3974)	Novel Protein sim. GBank gi488443(jb AA031319.1) [AF14457] - (AF144573) Mx-interacting protein kinase PKM [Mesocricetus auratus]	UNCLASSIFIED	22278999, 29331830, 265007, 265018, 21900706, 39537182, 264628, 264693, 18108377, 284635, 60170394, 22279002
1988	94122108 (3975, 3976)	Novel Protein sim. GBank gi5701727(jb AA03074.1) - (AB024723) alpha-1,3-D-mannoside beta-1,4-N-acetylglucosaminyltransferase IV homologue [Homo sapiens]	UNCLASSIFIED	264905, 264906, 264907, 264908, 264909, 264910, 264991, 284593, 264758, 264784, 264686, 264768, 265021, 264692, 264628, 264629, 39595955, 264630, 264635, 264636, 264837, 264838, 264839, 264463
1989	9125225 (3977, 3978)	Novel Protein sim. GBank gi5701727(jb AA03074.1) - (AB024723) alpha-1,3-D-mannoside beta-1,4-N-acetylglucosaminyltransferase IV homologue [Homo sapiens]	tubulin	60432045, 60432289, 32644045, 36182415, 264112, 265007, 3357402, 32644228, 18108385, 18108386, 264908, 39537182, 3957023, 393967, 394857109, 18108370, 22279000, 22279002
1990	8569888 (3979, 3980)	Novel Protein sim. GBank gi5701727(jb AA03074.1) - (AB024723) alpha-1,3-D-mannoside beta-1,4-N-acetylglucosaminyltransferase IV homologue [Homo sapiens]	UNCLASSIFIED	264508, 264757, 264764, 18108331
1991	56333114 (3981, 3982)	Novel Protein sim. GBank gi4240287(jb JAA74922.1) - (AB020706) KIAA0899 protein [Homo sapiens]	glycoprotein	18108394, 56182575, 22278994, 35696286, 56994075, 22278997, 22278999, 29331822, 29331824, 29331825, 60432289, 29331828, 264508, 264906, 264907, 264908, 56182435, 264510, 265007, 21900754, 33109954, 87168474, 285017, 265018, 265019, 264762, 18108351, 264763, 264693, 264389, 264288, 264685, 264766, 264687, 264769, 21900765, 21900768, 21900769, 3561857, 265020, 265021, 265022, 265023, 39595955, 264630, 264635, 264636, 264837, 264838, 264839, 83373044, 264555, 264557, 264639, 83373044, 18108394, 87168518, 60432113, 22279000, 22279002, 264564, 264485
1992	95317232 (3983, 3984)	Novel Protein sim. GBank gi2746532 (U39872) - ORF 73, contains large complex repeat CR 73 [Kaposi's sarcoma-associated herpesvirus]	sluot	29331827, 264506, 264907, 264909, 265007, 264603, 264766, 264688, 264788, 21900768, 264628, 264635, 264636, 18108385, 5626486, 264566, 264587
1993	80047463 (3985, 3986)	Novel Protein sim. GBank gi2954091 (U80761) - C1G28 alternate open reading frame [Homo sapiens]	UNCLASSIFIED	264592, 35696423

1194	94329114 (3087, 3088) gi3630077 gnl AA043622 - (AC00601) - (AC00601) similar to ACR; similar to AC031735 (P09232829) [Homo sapiens]	Contains protein domain (PF00585) SET domain	264448, 22278997, 22278998, 22278999, 22331822, 29331824, 29331825, 29331826, 29331827, 29331828, 264905, 264907, 264908, 265006, 60432226, 33657402, 33657403, 264737, 60433356, 264750, 33103936, 264751, 18108370, 264752, 264753, 264854, 264389, 264685, 264686, 264789, 21906765, 21906767, 21906768, 21906769, 265020, 264690, 18108382, 264493, 65274620, 18108370, 264635, 264555, 264556, 264557, 56182323, 83373044, 56529486, 22279000, 22279002, 264564
1195	95414353 (3989, 3990) gi4827040 ref NF_005110.1 pTRAP - thyroid hormone receptor-associated protein, 150 kDa subunit	UNCLASSIFIED	264448, 18108396, 22278994, 55994075, 22278996, 35696285, 22278997, 22278998, 264259, 29147620, 56182181, 29331824, 60432289, 29331826, 29331827, 35696052, 29146499, 264903, 264907, 66712502, 56182435, 265003, 265007, 265008, 265009, 60437735, 60433356, 33657402, 264595, 59812038, 33657084, 59811386, 65658542, 265010, 265011, 265012, 265013, 265014, 264751, 264752, 264448, 264783, 264784, 264238, 264785, 264686, 264789, 264789, 56181562, 264689, 21906765, 21906766, 21906767, 29148627, 21906768, 21906769, 29148629, 29148784, 265020, 265021, 264690, 18108361, 264693, 27486282, 27486284, 27486285, 18108370, 60431528, 18108374, 18108377, 35696423, 55811576, 65274791, 35696585, 264631, 264634, 264635, 264555, 264636, 60431850, 264557, 264558, 264659, 83373044, 20798451, 87169519, 264404, 59432113, 264567, 264564
1196	86234116 (3991, 3992) Novel Protein sim. GBank gi791146 emuc AA06020 - (K066020) - (K066020) - protein [Vigna unguiculata]	UNCLASSIFIED	
1197	87028423 (3993, 3994) Novel Protein sim. GBank gi45953 Hsp90A - protein phosphatase M130, cytosolic isoform [Oryzias latipes]	Contains protein domain (PF00023) Ark repeat	264908, 264909, 264967, 264933
1198	85262704 (3995, 3996) Novel Protein sim. GBank gi45953 Hsp90A - protein phosphatase M130, cytosolic isoform [Oryzias latipes]	UNCLASSIFIED	264113, 264685, 264555, 264567

1999	94324903 (3997, 3998)	Novel Protein sim. GBank gi15252512 gb AA04846.1 (AF072441) calcineurin binding protein cabn 1 [Homo sapiens]	Contains protein domain (PF00515) TPR Domain	UNCLASSIFIED	18108394, 18108397, 35696396, 60424260, 26531827, 35393378, 35696052, 35696052, 264512, 55611385, 265010, 265018, 265019, 59811150, 18108351, 264763, 264682, 264389, 264685, 264686, 56101692, 265020, 264691, 33657023, 264693, 33657109, 27486294, 18108370, 18108379, 35695855, 264634, 264635, 264636, 264555, 264557, 56182323, 18108382, 264559, 83373044, 60432113, 22279000, 264563, 264564, 264566
2000	96413705 (3999, 4000)	Novel Protein sim. GBank gi17232329 gi10153151 YA. SC.HPO. - HYPOTHETICAL 90.6 KD PROTEIN CID4.10 IN CHROMOSOME 1	UNCLASSIFIED	52646365, 52646842, 22278994, 22278995, 22278996, 35696286, 22278997, 22278998, 22278999, 264259, 52645000, 29331822, 29331826, 29331827, 29331828, 35696052, 264106, 29331830, 52644045, 265007, 265008, 60170831, 264592, 264593, 33657402, 60433438, 21906754, 52644286, 265017, 265018, 265019, 264761, 264389, 264288, 52644229, 21906765, 21906766, 21906767, 21906768, 21906769, 35695917, 265020, 265021, 60170515, 52644150, 33657023, 52674820, 52645129, 27486261, 27486282, 27486284, 35695763, 356960423, 35695955, 264631, 52644332, 56182323, 60170394, 83373044, 59526486, 22279002, 264566, 264567	
2001	95072534 (4001, 4002)	Novel Protein sim. GBank gi107560 pir B38637 - Ras inhibitor (clone JC265) - human (fragment)	UNCLASSIFIED	264789, 52644229, 52674572, 21906768, 22278996, 35696286, 35695917, 265020, 22278999, 264534, 264490, 264259, 264692, 60432289, 33657109, 35696052, 264508, 264509, 18108370, 60431528, 18108374, 356960423, 52674791, 35695955, 264510, 264511, 264512, 265009, 264634, 264636, 264555, 264556, 264638, 264557, 264558, 264559, 60433438, 83373044, 264759, 18108385, 265011, 264600, 264601, 60432113, 264603, 264604, 264605, 264448, 264288, 264785	
2002	80293368 (4003, 4004)	Novel Protein sim. GBank gi729433 pir 38657 JER60. BOVIN - PROBABLE PROTEIN DISULFIDE ISOMERASE ER-60 PRECURSOR (ERP60) (58 KD MICROSOMAL PROTEIN) (P59)	Contains protein domain (PF00085) thioredoxin	isomerase	264907, 265006, 264910, 264003, 264692, 264629, 18108374, 264556, 264557
2003	80074449 (4005, 4006)	Novel Protein sim. GBank gi85389 pir A27040 - neurofilament triplet N protein - chicken (fragment)	UNCLASSIFIED	264905, 264906, 264908, 264910, 264596, 265017, 18108351, 264692, 264629, 264634, 264655	

2004	95317318 (A007_4008)	Novel Protein sim. GBANK g0484249[emb]C4842320.11- (AL049966) hypothetical protein [homo sapiens]	Contains protein domain (PF00078): RNA recognition motif (A & B RRM, RBD, or RUP domain)	-dna_rna_bind	52645156, 52646842, 52946385, 55182575, 22278994, 22278995, 56994075, 22278996, 35662686, 22278997, 2278994, 22278996, 60433049, 254289, 52643500, 26331624, 25311620, 25331627, 35660052, 29331828, 35657070, 29331830, 264908, 264582, 60433356, 56568542, 23246317, 21506754, 253019, 264682, 264684, 9186599, 265017, 265018, 265019, 264765, 264683, 264288, 52644229, 21506755, 21506756, 21906767, 21906768, 21906769, 35668917, 265020, 265021, 265022, 52644150, 33957023, 52645128, 33857109, 33657182, 27486261, 21486262, 33857349, 27486265, 35959763, 18108374, 18108375, 18108377, 35964623, 35663855, 264631, 52644130, 264558, 18108385, 56526486, 87169518, 50432113, 264483, 264485
2005	87400864 (A009_4010)	Novel Protein sim. GBANK g0387950[emb]CAA87795[27812], similar to ubiquitin carboxyl-terminal hydrolase; cDNA EST EMBL:033368 comes from this gene; cDNA EST EMBL:033955 comes from this gene; cDNA EST EMBL:033822 comes from this gene; cDNA EST EMBL:034547 comes from this gene		ubiquitin	56182575, 55181686, 22278995, 22278996, 22278997, 22278998, 22278999, 22278995, 56182181, 60424265, 66714117, 35660502, 264906, 66712502, 264908, 264510, 60433356, 56568542, 256010, 265018, 265019, 264682, 264448, 264288, 264768, 29148627, 21906769, 29148784, 35663917, 60170615, 264691, 33657023, 66274620, 33857109, 55810764, 55811676, 36669885, 87169518, 55810763, 264483, 264482, 264488, 263694, 35660502, 264908, 264905, 264906, 264907, 264908, 264909, 264952, 264953, 264768, 265011, 264910, 264582, 264768, 264766, 264683, 264767, 264768, 264769, 55811957, 35663917, 265020, 264681, 264683, 264682, 264683, 264684, 264685, 264686, 264687, 264688, 264689, 264690, 264691, 264692, 264693, 264694, 264695, 264696, 264697, 264698, 264699, 264700, 264701, 264702, 264703, 264704, 264705, 264706, 264707, 264708, 264709, 264710, 264711, 264712, 264713, 264714, 264715, 264716, 264717, 264718, 264719, 264720, 264721, 264722, 264723, 264724, 264725, 264726, 264727, 264728, 264729, 264730, 264731, 264732, 264733, 264734, 264735, 264736, 264737, 264738, 264739, 264740, 264741, 264742, 264743, 264744, 264745, 264746, 264747, 264748, 264749, 264750, 264751, 264752, 264753, 264754, 264755, 264756, 264757, 264758, 264759, 264760, 264761, 264762, 264763, 264764, 264765, 264766, 264767, 264768, 264769, 264770, 264771, 264772, 264773, 264774, 264775, 264776, 264777, 264778, 264779, 264780, 264781, 264782, 264783, 264784, 264785, 264786, 264787, 264788, 264789, 264790, 264791, 264792, 264793, 264794, 264795, 264796, 264797, 264798, 264799, 264800, 264801, 264802, 264803, 264804, 264805, 264806, 264807, 264808, 264809, 264810, 264811, 264812, 264813, 264814, 264815, 264816, 264817, 264818, 264819, 264820, 264821, 264822, 264823, 264824, 264825, 264826, 264827, 264828, 264829, 264830, 264831, 264832, 264833, 264834, 264835, 264836, 264837, 264838, 264839, 264840, 264841, 264842, 264843, 264844, 264845, 264846, 264847, 264848, 264849, 264850, 264851, 264852, 264853, 264854, 264855, 264856, 264857, 264858, 264859, 264860, 264861, 264862, 264863, 264864, 264865, 264866, 264867, 264868, 264869, 264870, 264871, 264872, 264873, 264874, 264875, 264876, 264877, 264878, 264879, 264880, 264881, 264882, 264883, 264884, 264885, 264886, 264887, 264888, 264889, 264890, 264891, 264892, 264893, 264894, 264895, 264896, 264897, 264898, 264899, 264900, 264901, 264902, 264903, 264904, 264905, 264906, 264907, 264908, 264909, 264910, 264911, 264912, 264913, 264914, 264915, 264916, 264917, 264918, 264919, 264920, 264921, 264922, 264923, 264924, 264925, 264926, 264927, 264928, 264929, 264930, 264931, 264932, 264933, 264934, 264935, 264936, 264937, 264938, 264939, 264940, 264941, 264942, 264943, 264944, 264945, 264946, 264947, 264948, 264949, 264950, 264951, 264952, 264953, 264954, 264955, 264956, 264957, 264958, 264959, 264960, 264961, 264962, 264963, 264964, 264965, 264966, 264967, 264968, 264969, 264970, 264971, 264972, 264973, 264974, 264975, 264976, 264977, 264978, 264979, 264980, 264981, 264982, 264983, 264984, 264985, 264986, 264987, 264988, 264989, 264990, 264991, 264992, 264993, 264994, 2649

2009	85740240 (4017, 4018)	Novel Protein sim. GBank g 3822035 gb 3843435 12.1 - (AB018335) KIAA0792 protein [Homo sapiens]	UNCLASSIFIED	2277899, 264250, 264910, 264501, 265017, 264681, 264683, 21908768, 264691, 33857183, 33857349, 264631, 87168518, 264404, 22279002, 264583
2010	95422488 (4019, 4020)	Novel Protein sim. GBank g 5262629 emb CA945753.1 - (AL080164) hypothetical protein [Homo sapiens]	eph	52644507, 52645156, 52646842, 18108397, 65245472, 22278994, 56994475, 35696286, 22278996, 22278997, 22278999, 264259, 29331822, 52646080, 29331824, 29331825, 29331826, 29331827, 29331828, 264511, 265007, 264512, 265008, 265009, 60432229, 60433356, 21900754, 52646317, 33100994, 52644296, 87168474, 87168559, 265017, 265018, 265019, 264681, 264685, 264687, 52644229, 264688, 21907765, 21907766, 21907767, 21907768, 35695917, 33857183, 33857349, 264691, 264692, 33857023, 35696423, 65247291, 56994555, 18108378, 35696423, 65247291, 56994555, 264631, 264634, 60431850, 264637, 264638, 264631, 264634, 60431850, 264637, 264638, 62944332, 60170394, 18108385, 87168518, 22279002, 264584, 264585, 264586, 264587
2011	94328149 (4021, 4022)	Novel Protein sim. GBank g 3347953 AF076183 - cytosolic sorting protein PACS-1a [Rattus norvegicus]	UNCLASSIFIED	5618257, 56994075, 22278999, 264259, 29331824, 29331826, 29331827, 29331828, 35696052, 264906, 66712502, 265006, 265007, 265008, 265010, 265011, 265017, 265019, 264681, 264448, 264683, 264359, 264288, 264685, 264766, 264687, 21907765, 21907767, 21907768, 21907769, 265020, 265022, 264691, 33857023, 65247292, 33857100, 264429, 264537, 264539, 8337304, 87168518, 60432113, 22279002, 264584, 264585, 264586, 264587
2012	87772137 (4023, 4024)	Novel Protein sim. GBank g 1086578 U41020 - coded for by C. elegans cDNA Y41004.5, coded for by C. elegans cDNA Y41004.3, weekly similar to human SREBP-2 basic-helix-loop-helix-leucine zipper transcription factor [Caenorhabditis elegans]	Contains protein domain (PF00409) - Kinesin light chain repeat	8337304, 87168518, 60432113, 22279002, 264584, 264585, 264586, 264587, 264288, 264686, 21908764, 265020, 265022, 264288, 264686, 21908764, 265020, 265022, 33857023, 264693, 33857109, 55411576, 264632, 264555, 56182323, 264639, 18108383, 18108384, 18108388, 22279000, 22279002, 264667
2013	94643842 (4025, 4026)	Novel Protein sim. GBank g 4507985 pNP_003427.1p2NF1 - zinc finger protein 135 [clone p12.17]	Contains protein domain (PF00056) - Zinc finger, C2H2 type	18108399, 264908, 265007, 265010, 265016, 265019, 264680, 21906877, 265020, 264692, 265019, 264680, 21906877, 265020, 264692
2014	87347840 (4027, 4028)	Novel Protein sim. GBank g 412720 pP2093 MYO_HETR - MYELIN P0 PROTEIN PRECURSOR	UNCLASSIFIED	264488, 29331828, 264907, 264636, 264555, 264639, 264556

2015	86094922 (4020, 4030)	Novel Protein sim. GBank gi 82869 J52297 - etelensin - Vokox calien (fragment)	UNCLASSIFIED	56182573, 35698286, 264293, 35698052, 264508, 264906, 264907, 264510, 264512, 67160474, 265010, 264681, 264288, 264689, 264509, 264542, 35698655, 264659, 264553, 264554
2016	86298641 (4031, 4032)	Novel Protein sim. GBank gi 285046 pf S2641.3 - 1 complex protein Tcp-10 - mouse	stuct	264102, 264508, 264110, 265009, 33109954, 21906768, 265021, 33657109, 27486262, 263972, 18108374, 263976, 264555, 264564, 264685, 264638
2017	79464293 (4033, 4034)	Novel Protein sim. GBank gi 124735 pf 18175 INVQ_P1G - INVOLUCRIN	UNCLASSIFIED	264683
2018	79537087 (4035, 4036)	Novel Protein sim. GBank gi 2143910 pf J58216 - phosphatase-1 glycoen-binding (GJ)-chain - rat	phosphatase	264107, 264110, 264112, 265017, 263976
2019	87197900 (4037, 4038)	Novel Protein sim. GBank gi 2078483 (U43200) - ntlfreeze	UNCLASSIFIED	264259, 264508, 264391, 265018, 264682, 264288, 264686, 22279002
2020	94674476 (4039, 4040)	glycopeptide AFGP polypeptide precursor [Boreogadus tsalae]		
2021	98718818 (4041, 4042)	Novel Protein sim. GBank gi 15940700 pf GOL_PAT - ELONCATION		56994075, 264493, 33109954, 21906754, 21906768, 35657023, 33657109, 27486261, 67166518, 264509, 264542, 35698655, 264659, 264512
2022	96295665 (4043, 4044)	Novel Protein sim. GBank gi 2418956 (AC00135) - Parallel vclain storage protein (globulin-Bu) [Arabidopsis thaliana]		264517, 264707, 60170015, 18100385
2023	87722976 (4045, 4046)	Novel Protein sim. GBank gi 5410230 pf A042592.1 AF07334 - (AF073344) ulequin	Contains protein domain (PF00442) Ulequin carboxyl-terminal hydrolases family 2	18100394, 22278999, 264259, 264905, 264906, 264908, 264955, 264782, 264769, 264634, 264636, 67166518, 60432113, 22279000, 264482, 264565
2024	87896443 (4047, 4048)	specific protease 3 [Homo sapiens]		60433438, 265017, 264686, 264652, 264693, 264636
2025	87859863 (4049, 4050)	Novel Protein sim. GBank gi 475198 pf A020505.1 AC00701 - (AC007018) unknown protein [Arabidopsis thaliana]	UNCLASSIFIED	22278997, 2644509, 264906, 264909, 55412034, 265017, 265021, 265022, 60170615, 264556
2026	94122114 (4051, 4052)	Novel Protein sim. GBank gi 1055699 emb CAA69321 - (Y07752) pteroplin-S [Volvox carlini]	UNCLASSIFIED	56994075, 60432049, 264508, 66712502, 264112, 60170331, 87166559, 264288, 264688, 264685, 21906768, 33657109, 18100370, 264635, 18106385, 60432113, 22279000, 22279002, 264564, 264566, 264683
2027	80249001 (4053, 4054)		UNCLASSIFIED	263976, 264634, 264488

2028	94589844 (405, 406)	Novel Protein sim. GBank g1330245 (L35725) - coded for by C. elegans cDNA yk34b1.5, coded for by C. elegans cDNA yk31h10.5, coded for by C. elegans cDNA yk46a5.5, coded for by C. elegans cDNA yk46d5.5, coded for by C. elegans cDNA yk43c2.5, coded for by C. elegans cDNA yk46a8....	Contains protein domain (PF00632) - HECT-domain (ubiquitin-transferase).	ubiquitin	52644507, 52645156, 52646842, 56182755, 56994075, 56996062, 56997897, 56999407, 60432046, 264258, 52645080, 29331822, 29331824, 60714117, 29331825, 29331826, 29331827, 29331828, 35696062, 264696, 264697, 264698, 29331830, 52644045, 56182435, 265006, 265009, 60432229, 33657402, 264595, 264757, 55812038, 21906754, 52646317, 52644396, 265010, 265011, 87165559, 265017, 265018, 265019, 264448, 18103354, 264288, 264369, 264766, 52644228, 21906765, 21906768, 21906769, 21906769, 55811537, 21906767, 21906768, 21906769, 55811537, 21906767, 21906768, 21906769, 55811537, 52644159, 33657023, 65716261, 26463109, 52645125, 18103368, 27466261, 26463109, 27460264, 27466265, 35695763, 264632, 55811576, 35696423, 35695855, 264635, 264636, 52644332, 264558, 83373044, 55826485, 22279000, 22279002, 264563, 265009, 264595, 55638542, 264555, 264556, 264557, 264558, 264559, 83373044
2029	95362022 (4057, 4058)	Novel Protein sim. GBank g1589940 (AF073988) - Incidental Dysplasia protein 2 [Homo sapiens]	Contains protein domain (PF00821) - RhoGEF domain	UNCLASSIFIED	264557, 264558, 264559, 83373044
2030	91213734 (4059, 4060)	Novel Protein sim. GBank g160361040-1823 (JAC00439 - (AC004850) similar to H4B1; similar to B0024386 (P102783-06) [Homo sapiens])	Contains protein domain (PF00066) - Zinc finger, C2H2 type	transcript factor	264259, 29331822, 29331824, 29331825, 60714117, 60432289, 29331826, 264108, 6071502, 294278, 265009, 265018, 265019, 264681, 264682, 264684, 264685, 56181862, 264689, 21906769, 265022, 264692, 264693, 264694, 264695, 18103368, 264694, 264636, 264558, 18103385, 87160516, 22279002, 264565
2031	8025261 (4061, 4062)				264561, 55811567, 18103395, 264557, 264558, 18103392, 18103394
2032	91232607 (4063, 4064)	Novel Protein sim. GBank g15885151 (db BA03033.1) - (AB029000) KIAA1077 protein [Homo sapiens]	Contains protein domain (PF00884) - Sulfatase	hydrolase	55774572, 35698286, 29331824, 264908, 265006, 264953, 265018, 264288, 264886, 264769, 21906766, 21906767, 29148627, 264628, 35696423, 264634, 264556, 18103381, 60710394, 264559, 83373044, 18103385, 264482, 264484
2033	95008009 (4065, 4066)	Novel Protein sim. GBank g160361040-1823 (JAC00439 - (AC004850) similar to H4B1; similar to B0024386 (P102783-06) [Homo sapiens])		synthase	55811562, 264628, 264632, 264555, 264558
2034	91232529 (4067, 4068)	Novel Protein sim. GBank g14829884 (pepNP_005117, IIPROD1 - UNKNOWN)	Contains protein domain (PF00078) - RNA recognition motif (a.k.a. RRM, RBD, or RNP domain)	dna_rna_bind	22276956, 22276959, 264607, 79331330, 264608, 264609, 264610, 264611, 264612, 264613, 21906767, 21906768, 21906769, 35697108, 83373044, 55825646

2035	03563451 (4/89, 4/07)			264389, 264686, 265027, 2656486, 264567
2036	07115833 (4/01, 4/02)			26331827, 26331238, 264682, 264389, 2614867, 60432113
2037	94324833 (4/03, 4/04)	Novel Protein sim. GBank, g12734081 (4/200195) - similar to oxytetracycline-binding proteins [Caenorhabditis elegans]	UNCLASSIFIED	35274572, 22278995, 22278996, 265094075, 3569266, 22278997, 22278998, 22278999, 26331824, 60432289, 26331286, 26331825, 264686, 26501800, 60712502, 26482435, 265008, 265009, 60170831, 264594, 5812038, 33105954, 21906754, 87168559, 265017, 265018, 265019, 264762, 264369, 264388, 21902765, 21906767, 21906768, 21906789, 55811857, 35693917, 265020, 265021, 265022, 59641150, 33657023, 33657109, 33657182, 35692763, 35695855, 264632, 264834, 264636, 59182323, 83373044, 60432113, 22279000, 22279002, 264563
2038	95422384 (4/03, 4/06)	Novel Protein sim. GBank, g1380826jpmh(CA607858) - (233785) predicted using GeneRuler, similar to RNA recognition motif (aka RRM, RBD, or RNP domain); cDNA EST EMBL M75823 comes from this gene; cDNA EST EMBL M75823 comes from this gene; cDNA EST EMBL D27559 comes from this gene...	UNCLASSIFIED	22278995, 22278996, 59994075, 264539, 29331824, 35696052, 264905, 264906, 52644045, 265007, 265009, 87168559, 265017, 18108351, 264448, 264389, 264786, 264787, 264686, 18108358, 21906765, 21906766, 21906767, 264686, 264687, 18108359, 33657109, 21906765, 18108362, 33657109, 21906765, 264686, 18108374, 18108379, 35696423, 65274791, 264632, 264636, 18108383, 83373044, 18108385, 87168518, 22279000, 22279002, 264563, 264564, 264566
2039	95514626 (4/07, 4/08)	Novel Protein sim. GBank, g1224653jdj(BA023813) - (40023541) KIAA0356 [Homo sapiens]	UNCLASSIFIED	22278997, 264259, 29331822, 264905, 264908, 264907, 264908, 264909, 264510, 265009, 264910, 264593, 264758, 265011, 265018, 264762, 264288, 264766, 264768, 264769, 21906766, 33657023, 264692, 264693, 33657109, 35696423, 264631, 264632, 264634, 264635, 264636, 264637, 264639, 87168519, 264498
2040	95338417 (4/07, 4/08)	Novel Protein sim. GBank, g1250825isp(P70700)P2A, MOUSE - DNA DIRECTED RNA POLYMERASE 1135 KD POLYPEPTIDE (RNA POLYMERASE I SUBUNIT 2) (RPA135)	UNCLASSIFIED	264592, 264688, 22278998, 35696082, 264906, 264907, 264910, 265018, 264605, 265019, 18108351, 264758, 264762, 21906765, 265021, 265022, 264692, 33657109, 264629, 264630, 35696423, 35696855, 264637, 264638, 264639, 264640, 264641, 264642, 264643, 264644, 264645, 264646, 264647, 264648, 264649, 264650, 264651, 264652, 264653, 264654, 264655, 264656

[illegible]

2050	78433035 (4089, 4100)				264683	264468, 264258, 264508, 264906, 264907, 264769, 18108374, 3068443, 264563,
2051	87760168 (4101, 4102)				UNCLASSIFIED	
2052	86090393 (4103, 4104)	Novel Protein sim. GBank g1452883gip4A021812.11 - (AF1341726) G9A (Homo sapiens)	Contains protein domain (PF00856) - kinase SET domain		UNCLASSIFIED	264489, 264509, 15690652, 264508, 264905, 264489, 264509, 264907, 264908, 264909, 264513, 264511, 265009, 264910, 60170831, 264592, 264758, 265011, 264605, 264768, 264682, 264764, 264369, 264766, 264688, 264768, 264769, 50644229, 264689, 35695917, 33657023, 33657109, 264628, 18108374, 35696423, 55811576, 35695855, 264630, 264631, 264632, 264634, 264635, 264636, 264638, 264639, 18108385, 56526486, 60432113, 264563, 264654, 264586, 264488, 264567, 264488, 263994, 35990052, 264508, 264905, 264509, 264906, 264907, 264508, 264909, 264113, 264511, 265009, 264910, 60170831, 264592, 264758, 264688, 264768, 264769, 50644229, 264689, 35695917, 33657023, 33657109, 264628, 18108374, 35696423, 264630, 264631, 264632, 264634, 264635, 264636, 264556, 264638, 264639, 18108385, 56526486, 60432113, 264563, 264584, 264586, 264488, 264567
2053	87763078 (4105, 4106)	Novel Protein sim. GBank g12995448emb(CAA75113) - (Y14846) mydline 1 protein (Mus musculus)			UNCLASSIFIED	22778986, 22778987, 264259, 78331822, 264102, 264508, 35695917, 263972, 264482
2054	95359537 (4107, 4108)	Novel Protein sim. GBank g1376262emb(CAB02090) - (Z79754) similar to C2 domain [Caenorhabditis elegans]	Contains protein domain (PF00168) - C2 domain		UNCLASSIFIED	60424178, 264094, 264259, 29331825, 60424269, 264096, 60432229, 60433356, 87166559, 265019, 264780, 264288, 264686, 11906769, 33657023, 264993, 55810764, 55811576, 264535, 58162323, 60432113, 264508, 264905, 264906, 264907, 264908, 264909, 264513, 264511, 265009, 264910, 60170831, 264592, 264758, 264688, 264768, 264769, 50644229, 264689, 35695917, 33657023, 33657109, 264628, 18108374, 35696423, 264630, 264631, 264632, 264634, 264635, 264636, 264556, 264638, 264639, 18108385, 56526486, 60432113, 264563, 264584, 264586, 264488, 264567
2055	86259449 (4109, 4110)	Novel Protein sim. GBank g16353748bIMAD2268.1AF15913 - (AF159133) SIR2-like protein (Oryza sativa subsp. indica)			UNCLASSIFIED	264489, 264509, 15690652, 264508, 264905, 264489, 264509, 264907, 264908, 264909, 264513, 264511, 265009, 264910, 60170831, 264592, 264758, 264688, 264768, 264769, 50644229, 264689, 35695917, 33657023, 33657109, 264628, 18108374, 35696423, 264630, 264631, 264632, 264634, 264635, 264636, 264556, 264638, 264639, 18108385, 56526486, 60432113, 264563, 264584, 264586, 264488, 264567

2056	(88177596 (4111, 4112))	Novel Protein sim. GBank g14452650(meNP_005042_1)QARS - glutamine-tRNA synthetase	Contains protein domain (PF00749) tRNA synthetase class I (E and Q)	synthase	264448, 524515, 561925, 2227894, 35660286, 56984075, 2227896, 2227898, 2227899, 50432049, 60432289, 29331827, 25331828, 264104, 264906, 264908, 265006, 265008, 265010, 264981, 60432228, 60433438, 18103348, 21906754, 33857094, 52644296, 87168474, 265010, 87185559, 265017, 265018, 264760, 18108351, 264681, 264682, 264448, 264683, 264366, 264369, 264388, 264685, 264687, 264688, 264689, 21906765, 21906766, 21906767, 21906769, 55811957, 35695917, 265022, 33657023, 18108362, 33657108, 18108368, 33657182, 27486261, 27486264, 27486265, 33657349, 264626, 18108370, 26462611, 18108371, 18108372, 35696123, 55811576, 20381152, 264636, 264892, 18108385, 18108388, 18108398, 87188518, 264482, 264565, 264566, 264567, 52646842, 52646865, 56182575, 35696288, 22278986, 22278987, 22278989, 2645083, 52645080, 35696052, 29331828, 33656870, 265009, 52646317, 55811386, 52644286, 52644229, 21906769, 35695917, 265021, 60170615, 52644150, 33657109, 33657182, 27486261, 27486262, 35695763, 35696423, 35696555, 52644332, 265007, 265008, 265091, 28331823, 264682, 264686, 264691, 264693, 22279002
2057	(87877605 (4113, 4114))	Novel Protein sim. GBank g1728850jpp08640AMYH_ YEAST - GLUCOAMYLASE S1S2 PRECURSOR (GLUCAN 1,4-ALPHA- GLUCOSIDASE) (1,4-ALPHA-D-GLUCAN GLUCOHYDROLASE)		UNCLASSIFIED	52645080, 35696052, 29331828, 33656870, 265009, 52646317, 55811386, 52644286, 52644229, 21906769, 35695917, 265021, 60170615, 52644150, 33657109, 33657182, 27486261, 27486262, 35695763, 35696423, 35696555, 52644332, 265007, 265008, 265091, 28331823, 264682, 264686, 264691, 264693, 22279002
2058	(8276596 (4115, 4116))	Novel Protein sim. GBank g111974jpp13963JEXTN_TDBAC - EXTENSIN PRECURSOR (CELL WALL HYDROXYPROLINE-RICH GLYCOPROTEIN)		UNCLASSIFIED	56182575, 29331824, 26318326, 264910, 55811957, 18108370, 18108371, 18108372
2060	(30505600 (4119, 4120))	Novel Protein sim. GBank g12811122 (J87318) - NADOC 2 (Kempner levels)		UNCLASSIFIED	56182575, 29331824, 26318326, 264910, 55811957, 18108370, 18108371, 18108372

2075	9431486 (4149, 4150)	Novel Protein sm. GBank g 5138520 g AAQ0382.1 - (AF050680) transcription factor IIB (Homo sapiens)			18103394, 22278994, 22278996, 35696286, 22278998, 22278999, 264259, 29331822, 29331825, 29331827, 35696052, 29331828, 294950, 264907, 264908, 294950, 294951, 294952, 294953, 294954, 294955, 294956, 294957, 294958, 294959, 294960, 294961, 294962, 294963, 294964, 294965, 294966, 294967, 294968, 294969, 294970, 294971, 294972, 294973, 294974, 294975, 294976, 294977, 294978, 294979, 294980, 294981, 294982, 294983, 294984, 294985, 294986, 294987, 294988, 294989, 294990, 294991, 294992, 294993, 294994, 294995, 294996, 294997, 294998, 294999, 295000, 295001, 295002, 295003, 295004, 295005, 295006, 295007, 295008, 295009, 295010, 295011, 295012, 295013, 295014, 295015, 295016, 295017, 295018, 295019, 295020, 295021, 295022, 295023, 295024, 295025, 295026, 295027, 295028, 295029, 295030, 295031, 295032, 295033, 295034, 295035, 295036, 295037, 295038, 295039, 295040, 295041, 295042, 295043, 295044, 295045, 295046, 295047, 295048, 295049, 295050, 295051, 295052, 295053, 295054, 295055, 295056, 295057, 295058, 295059, 295060, 295061, 295062, 295063, 295064, 295065, 295066, 295067, 295068, 295069, 295070, 295071, 295072, 295073, 295074, 295075, 295076, 295077, 295078, 295079, 295080, 295081, 295082, 295083, 295084, 295085, 295086, 295087, 295088, 295089, 295090, 295091, 295092, 295093, 295094, 295095, 295096, 295097, 295098, 295099, 295100, 295101, 295102, 295103, 295104, 295105, 295106, 295107, 295108, 295109, 295110, 295111, 295112, 295113, 295114, 295115, 295116, 295117, 295118, 295119, 295120, 295121, 295122, 295123, 295124, 295125, 295126, 295127, 295128, 295129, 295130, 295131, 295132, 295133, 295134, 295135, 295136, 295137, 295138, 295139, 295140, 295141, 295142, 295143, 295144, 295145, 295146, 295147, 295148, 295149, 295150, 295151, 295152, 295153, 295154, 295155, 295156, 295157, 295158, 295159, 295160, 295161, 295162, 295163, 295164, 295165, 295166, 295167, 295168, 295169, 295170, 295171, 295172, 295173, 295174, 295175, 295176, 295177, 295178, 295179, 295180, 295181, 295182, 295183, 295184, 295185, 295186, 295187, 295188, 295189, 295190, 295191, 295192, 295193, 295194, 295195, 295196, 295197, 295198, 295199, 295200, 295201, 295202, 295203, 295204, 295205, 295206, 295207, 295208, 295209, 295210, 295211, 295212, 295213, 295214, 295215, 295216, 295217, 295218, 295219, 295220, 295221, 295222, 295223, 295224, 295225, 295226, 295227, 295228, 295229, 295230, 295231, 295232, 295233, 295234, 295235, 295236, 295237, 295238, 295239, 295240, 295241, 295242, 295243, 295244, 295245, 295246, 295247, 295248, 295249, 295250, 295251, 295252, 295253, 295254, 295255, 295256, 295257, 295258, 295259, 295260, 295261, 295262, 295263, 295264, 295265, 295266, 295267, 295268, 295269, 295270, 295271, 295272, 295273, 295274, 295275, 295276, 295277, 295278, 295279, 295280, 295281, 295282, 295283, 295284, 295285, 295286, 295287, 295288, 295289, 295290, 295291, 295292, 295293, 295294, 295295, 295296, 295297, 295298, 295299, 295300, 295301, 295302, 295303, 295304, 295305, 295306, 295307, 295308, 295309, 295310, 295311, 295312, 295313, 295314, 295315, 295316, 295317, 295318, 295319, 295320, 295321, 295322, 295323, 295324, 295325, 295326, 295327, 295328, 295329, 295330, 295331, 295332, 295333, 295334, 295335, 295336, 295337, 295338, 295339, 295340, 295341, 295342, 295343, 295344, 295345, 295346, 295347, 295348, 295349, 295350, 295351, 295352, 295353, 295354, 295355, 295356, 295357, 295358, 295359, 295360, 295361, 295362, 295363, 295364, 295365, 295366, 295367, 295368, 295369, 295370, 295371, 295372, 295373, 295374, 295375, 295376, 295377, 295378, 295379, 295380, 295381, 295382, 295383, 295384, 295385, 295386, 295387, 295388, 295389, 295390, 295391, 295392, 295393, 295394, 295395, 295396, 295397, 295398, 295399, 295400, 295401, 295402, 295403, 295404, 295405, 295406, 295407, 295408, 295409, 295410, 295411, 295412, 295413, 295414, 295415, 295416, 295417, 295418, 295419, 295420, 295421, 295422, 295423, 295424, 295425, 295426, 295427, 295428, 295429, 295430, 295
------	----------------------	---	--	--	---

2000	88222470 (4178, 4180)				22278995, 22278998, 22278999, 264295, 29331626, 35696052, 384910, 39357402, 60433493, 33109954, 87168474, 87168599, 295016, 265019, 264681, 264684, 264696, 295016, 265019, 264681, 264684, 264696, 21906765, 21906767, 21906769, 35696817, 21906766, 21906768, 21906769, 35696842, 265022, 60170615, 33657023, 35698423, 35698585, 264952, 18103537, 22279000, 2653994, 264905, 264908, 264511, 264312, 265008, 264910, 55811398, 264768, 264768, 56181562, 21906765, 21906768, 21906769, 265022, 264628, 264563, 264567, 265022, 264628, 264563, 264567
2081	95209181 (4181, 4182)	Novel Protein sim. GBank gH4560997 gb AAC24571.1 AF12108 - (AF12108) cAMP inducible 2 protein [Mus musculus]	UNCLASSIFIED		22278997, 22278999, 68712502, 87168559, 284683, 265021, 264486
2092	98223605 (4183, 4184)	Novel Protein sim. GBank gH4560997 gb AAC24571.1 AF12108 - (AF12108) cAMP inducible 2 protein [Mus musculus]	homeobox		22278997, 22278999, 68712502, 87168559, 284683, 265021, 264486
2093	87405073 (4185, 4186)	Novel Protein sim. GBank gH4560997 gb AAC24571.1 AF12108 - (AF12108) cAMP inducible 2 protein [Mus musculus]	Contains protein domain (PF00013) - Mitochondrial carrier proteins		22278997, 22278999, 68712502, 87168559, 284683, 265021, 264486
2094	91230929 (4187, 4188)	Novel Protein sim. GBank gH4560997 gb AAC24571.1 AF12108 - (AF12108) cAMP inducible 2 protein [Mus musculus]	homeobox		22278997, 22278999, 68712502, 87168559, 284683, 265021, 264486
2095	95351526 (4189, 4190)	Novel Protein sim. GBank gH4560997 gb AAC24571.1 AF12108 - (AF12108) cAMP inducible 2 protein [Mus musculus]	homeobox		22278997, 22278999, 68712502, 87168559, 284683, 265021, 264486
2096	94119760 (4191, 4192)	Novel Protein sim. GBank gH4560997 gb AAC24571.1 AF12108 - (AF12108) cAMP inducible 2 protein [Mus musculus]	homeobox		22278997, 22278999, 68712502, 87168559, 284683, 265021, 264486

2097	95322772 (4153, 4194)	Novel Protein sim. GBank gH5174501[neNP_000051.1]pLYF1 - zinc finger protein. subfamily 1A, 1 (karcos)	Contains protein domain (PF00046) - Zinc finger, C2H2 type	transfer factor	65274572, 264511, 265010, 264500, 265017, 264448, 264288, 265021, 30170015, 264692, 32581705, 18108370, 264335, 264463, 30894073, 264236, 264286, 265020, 264563
2098	87760340 (4195, 4196)	Novel Protein sim. GBank gH475820[neNP_004081.1]DUSP - dual specificity phosphatase 3 (vaccinia virus phosphatase VH1-related)	Contains protein domain (PF00742) - Dual specificity phosphatase, catalytic domain	- phosphatase	
2099	95412927 (4197, 4198)	Novel Protein sim. GBank gH639559 [AF026554] - pyruvate dehydrogenase phosphatase regulatory subunit precursor, PDPT [Bos taurus]	Contains protein domain (PF00025) - ADP-ribosylation factor family	phosphatase	65274572, 264905, 35274444, 264691, 264636, 264555
2100	95332656 (4199, 4200)	Novel Protein sim. GBank gH38116[emj]CA1816514 - (A02811) similar to ADP-ribosylation factor; cDNA EST EMBL:080337 comes from this gene; cDNA EST EMBL:098826 comes from this gene; cDNA EST y4291b4.5 comes from this gene; cDNA EST y4...	Contains protein domain (PF00025) - ADP-ribosylation factor family	- nucl. reprot	56162575, 22276965, 22276966, 22276967, 22276968, 60432040, 264259, 22331832, 29331824, 29331825, 29331827, 29331828, 29146498, 264909, 265008, 295009, 284910, 264591, 60432229, 60433356, 33857402, 264758, 21900754, 36658442, 87168474, 265017, 265018, 265019, 264681, 18108351, 264762, 264448, 264359, 264281, 18108355, 21900759, 21900765, 21900767, 21900768, 21900769, 21900770, 21900771, 21900772, 18108374, 32666153, 359271, 3857023, 87168518, 60432113, 22276966, 22276967, 18108390, 60432113, 22276966, 22276967, 264091, 29331824, 364705, 265007, 265010, 18108390
2101	87762904 (4201, 4202)	Novel Protein sim. GBank gH458946[gbj]BA476761.1] - (A8012808) mBOCT [Mus musculus]	UNCLASSIFIED	UNCLASSIFIED	264448, 264489, 35665268, 264259, 35665032, 264508, 264605, 264907, 264908, 264909, 264511, 264512, 264591, 264593, 60433356, 264758, 264601, 264605, 264790, 18108351, 264448, 264764, 264288, 264767, 264768, 21900759, 35665917, 18108374, 264634, 264555, 264566, 264593, 264482, 264486
2102	87770461 (4203, 4204)	Novel Protein sim. GBank gH38714[gmj]CA67423.1] - (Z73100) predicted using GeneFinder [Caenorhabditis elegans]	UNCLASSIFIED	UNCLASSIFIED	
2103	95413576 (4205, 4206)	Novel Protein sim. GBank gH4240159[gbj]BA474858.1] - (A8020942) KIAA0835 protein [homo sapiens]	Contains protein domain (PF01530) - Zinc finger, C2H2 type	- transcript factor	65274572, 59894075, 22276959, 264259, 29331824, 29331825, 35666002, 29331828, 60712502, 265009, 30170031, 264395, 35194955, 35638542, 87165599, 265017, 265018, 264762, 21900759, 21900768, 265027, 33857023, 27448262, 35665917, 35665918, 35665919, 60431526, 18108374, 55511676, 56182323, 18108387, 87168518, 60432113, 264484
2104	85776161 (4207, 4208)		UNCLASSIFIED	UNCLASSIFIED	264592, 264604, 22276900

2105	94648030 (4209, 4210)	Novel Protein sim. GBank g1170332 (U8045) - coded for by C. elegans cDNA Y13c4.3; coded for by C. elegans cDNA Y621p6 3; coded for by C. elegans cDNA CEMSE1BF coded for by C. elegans cDNA Y126b1.3; coded for by C. elegans cDNA Y65h18.3; coded for by C. elegans cDNA Y65h8 ...	UNCLASSIFIED	26448, 26192176, 22978904, 26564071, 22757898, 22218007, 22748999, 22222896, 60432044, 264459, 20331822, 20331824, 20331825, 60432280, 20331827, 20331828, 264508, 264905, 264509, 264907, 20331830, 52844045, 264510, 264511, 264907, 264512, 265009, 60170831, 60432229, 33557402, 60433395, 264555, 60433438, 264758, 33557084, 87168474, 265010, 87168559, 265017, 265018, 265019, 264782, 18108351, 264684, 18108354, 264288, 264686, 52544229, 18108359, 21906765, 21906766, 21906767, 21906768, 21906769, 33695917, 265020, 265022, 60170515, 52544150, 60431528, 18108374, 65274791, 33695955, 264633, 60170394, 264639, 264598, 18108387, 36526466, 87168518, 60432113, 264594, 264595, 264597, 265006, 265016, 264596, 264597, 264906, 264639
2106	83365475 (4211, 4212)	Novel Protein sim. GBank g1381524 (emb)(CA43883) - (Z70038), ZK1067 4 (Caenorhabditis elegans)	UNCLASSIFIED	264903, 264906, 264907, 264908, 264909, 264758, 265011, 264900, 264801, 264754, 264766, 264767, 264768, 264769, 264693, 264679, 33595555, 264632, 264634, 264635, 264638, 264639, 83373044, 264486, 18108345, 264769, 18108370, 18108374, 264555, 264556, 264457, 264458, 264564
2107	78527662 (4213, 4214)	Novel Protein sim. GBank g1375688 (A000367) - Contains similarity to ubiquitin carboxyl-terminal hydrolase 14 gp(235927 from S. cerevisiae [Arabidopsis thaliana])	UNCLASSIFIED	264903, 264906, 264907, 264908, 264909, 264758, 265011, 264900, 264801, 264754, 264766, 264767, 264768, 264769, 264693, 264679, 33595555, 264632, 264634, 264635, 264638, 264639, 83373044, 264486, 18108345, 264769, 18108370, 18108374, 264555, 264556, 264457, 264458, 264564
2108	80478719 (4217, 4218)	Novel Protein sim. GBank g1451043 (p157871 - bat2 protein - human)	UNCLASSIFIED	264903, 264906, 264907, 264908, 264909, 264758, 265011, 264900, 264801, 264754, 264766, 264767, 264768, 264769, 264693, 264679, 33595555, 264632, 264634, 264635, 264638, 264639, 83373044, 264486, 18108345, 264769, 18108370, 18108374, 264555, 264556, 264457, 264458, 264564
2109	87759075 (4219, 4220)	Novel Protein sim. GBank g12143530 (p157871 - bat2 protein - human)	UNCLASSIFIED	264903, 264906, 264907, 264908, 264909, 264758, 265011, 264900, 264801, 264754, 264766, 264767, 264768, 264769, 264693, 264679, 33595555, 264632, 264634, 264635, 264638, 264639, 83373044, 264486, 18108345, 264769, 18108370, 18108374, 264555, 264556, 264457, 264458, 264564
2110	87818419 (4221, 4222)	Novel Protein sim. GBank g12143530 (p157871 - bat2 protein - human)	UNCLASSIFIED	264903, 264906, 264907, 264908, 264909, 264758, 265011, 264900, 264801, 264754, 264766, 264767, 264768, 264769, 264693, 264679, 33595555, 264632, 264634, 264635, 264638, 264639, 83373044, 264486, 18108345, 264769, 18108370, 18108374, 264555, 264556, 264457, 264458, 264564
2111	87523783 (4223, 4224)	Novel Protein sim. GBank g12143530 (p157871 - bat2 protein - human)	UNCLASSIFIED	264903, 264906, 264907, 264908, 264909, 264758, 265011, 264900, 264801, 264754, 264766, 264767, 264768, 264769, 264693, 264679, 33595555, 264632, 264634, 264635, 264638, 264639, 83373044, 264486, 18108345, 264769, 18108370, 18108374, 264555, 264556, 264457, 264458, 264564
2112	87523783 (4223, 4224)	Novel Protein sim. GBank g12143530 (p157871 - bat2 protein - human)	UNCLASSIFIED	264903, 264906, 264907, 264908, 264909, 264758, 265011, 264900, 264801, 264754, 264766, 264767, 264768, 264769, 264693, 264679, 33595555, 264632, 264634, 264635, 264638, 264639, 83373044, 264486, 18108345, 264769, 18108370, 18108374, 264555, 264556, 264457, 264458, 264564
2113	76841388 (4225, 4226)	Novel Protein sim. GBank g12143530 (p157871 - bat2 protein - human)	UNCLASSIFIED	264903, 264906, 264907, 264908, 264909, 264758, 265011, 264900, 264801, 264754, 264766, 264767, 264768, 264769, 264693, 264679, 33595555, 264632, 264634, 264635, 264638, 264639, 83373044, 264486, 18108345, 264769, 18108370, 18108374, 264555, 264556, 264457, 264458, 264564
2114	87889342 (4227, 4228)	Novel Protein sim. GBank g13327184 (db)(BA31890) - (AF100950) protocadherin (Rattus norvegicus)	UNCLASSIFIED	264903, 264906, 264907, 264908, 264909, 264758, 265011, 264900, 264801, 264754, 264766, 264767, 264768, 264769, 264693, 264679, 33595555, 264632, 264634, 264635, 264638, 264639, 83373044, 264486, 18108345, 264769, 18108370, 18108374, 264555, 264556, 264457, 264458, 264564
2115	909933785 (4229, 4230)	Novel Protein sim. GBank g14757890 (refNP_004328 1pC20R - chromosome 8 open reading frame 1)	UNCLASSIFIED	264903, 264906, 264907, 264908, 264909, 264758, 265011, 264900, 264801, 264754, 264766, 264767, 264768, 264769, 264693, 264679, 33595555, 264632, 264634, 264635, 264638, 264639, 83373044, 264486, 18108345, 264769, 18108370, 18108374, 264555, 264556, 264457, 264458, 264564

2127	81118652 (4253, 4254)	Novel Protein sim. GBank gi14688453bb AD31315.1 AF14323.3 (AF14323.3) apoptosis related protein AFR-2 [Homo sapiens]		35595052, 29331826, 35595052, 264508, 264509, 264905, 264908, 264907, 264908, 264909, 264510, 265005, 264511, 264512, 265007, 265009, 264910, 264758, 265011, 264600, 264601, 264604, 264762, 264763, 264766, 264637, 264768, 264769, 264689, 35595917, 264690, 264691, 264692, 264693, 264629, 18103374, 35595955, 264632, 264634, 264635, 264636, 264637, 264638, 264639, 18103385, 264663, 264664, 264657, 35182573, 35595952, 56182181, 29331824, 80432299, 35596052, 264905, 264907, 8671502, 264906, 264909, 264510, 264512, 265009, 264910, 264931, 55912038, 265016, 5591704, 264286, 264395, 264697, 264768, 264769, 264770, 264771, 264772, 264773, 264632, 264634, 264635, 264637, 56182323, 264639, 18103384, 18103388, 264663, 264667
2128	87414282 (4255, 4256)			UNCLASSIFIED
2129	95102089 (4257, 4258)			UNCLASSIFIED
2130	95417144 (4259, 4260)	Novel Protein sim. GBank gi2649255 (AE001012) - conserved hypothetical protein [Archaeoglobus fulgidus]		UNCLASSIFIED
2131	8572065 (4261, 4262)	Novel Protein sim. GBank gi106886 (U4176) - (Similar to potassium channel protein, [Cnemidobolus elegans])	Contains protein domain (PF00805) - Pentapeptide repeats (8 copies)	potassium_channel
2132	95381096 (4263, 4264)	Novel Protein sim. GBank gi509373 db BA045293.1 - (AB028944) KIAA1021 protein [Homo sapiens]	Contains protein domain (PF00122) - ET-EZ ATPase	ATPase_associated

2133	9535-1539 (4265, 4266)	Novel Protein sim. GBank #I420486 (AC006069) - hypothetical protein [Neosporus italicus]	UNCLASSIFIED	59424173, 52546335, 52546842, 56994075, 5956285, 2278997, 2278998, 50432049, 5956286, 2278997, 2278998, 50432049, 29331839, 35695952, 264605, 264608, 264607, 66712502, 36331830, 59182435, 265006, 264512, 265008, 60431735, 60433956, 33657542, 58112038, 33109954, 21906754, 55811386, 265010, 264803, 265017, 265018, 265019, 55811150, 18108351, 264682, 264389, 264285, 52644239, 56181562, 21906785, 21906786, 21906767, 21908788, 21906789, 55811957, 35695917, 265020, 265021, 60170615, 33657023, 33657109, 60431528, 18108374, 356959423, 65274791, 35695855, 264634, 60431850, 18108380, 56182323, 83373044, 18108385, 18108387, 60432113, 22729002, 264583, 264585
2134	95412597 (4267, 4268)	Novel Protein sim. GBank #J387535 [homoNC-I609415] - (Z96047) DY3.6 [C.annoniaute elegans]	UNCLASSIFIED	56181689, 356959286, 21906754, 55811388, 265011, 265017, 18108351, 264785, 264786, 33657023, 21906788, 35695917, 265020, 265021, 60170615, 33657023, 33657109, 60431528, 18108374, 35695942, 264634, 264635, 264636, 264637, 56182323, 264639, 22729002, 264584
2135	88079813 (4269, 4270)	Novel Protein sim. GBank #J56955596 [BAA31063.1] (A8029034) KIAA1111 protein [homo sapiens]	UNCLASSIFIED	22278399, 26331828, 35695952, 264605, 264608, 264910, 265009, 264591, 264758, 52646317, 265011, 87188559, 264601, 18108351, 264446, 264683, 264684, 264689, 18108359, 264691, 33657023, 264692, 35695763, 264629, 35695955, 264831, 264835, 264636, 264637, 56182323, 264639, 22729002, 264584
2136	84346479 (4271, 4272)	Novel Protein sim. GBank #J2682167 [BAA23715] - (A8077903) KIAA0443 [homo sapiens]	UNCLASSIFIED	2646369, 264910, 265009, 264591, 264758, 52646317, 265011, 87188559, 264601, 18108351, 264446, 264683, 264684, 264689, 18108359, 264691, 33657023, 264692, 35695763, 264629, 35695955, 264831, 264835, 264636, 264637, 56182323, 264639, 22729002, 264584
2137	87537718 (4273, 4274)	Novel Protein sim. GBank #J4884110 [homoCA43262.1] - (AL050080) hypothetical protein [homo sapiens]	UNCLASSIFIED	2646369, 264910, 265009, 264591, 264758, 52646317, 265011, 87188559, 264601, 18108351, 264446, 264683, 264684, 264689, 18108359, 264691, 33657023, 264692, 35695763, 264629, 35695955, 264831, 264835, 264636, 264637, 56182323, 264639, 22729002, 264584
2138	87356446 (4275, 4276)	Novel Protein sim. GBank #J5747939 [AC006068.1] - (U87604) 50 kDa protein [Calobacter crescentius]	ATPase-associated	2646369, 264910, 265009, 264591, 264758, 52646317, 265011, 87188559, 264601, 18108351, 264446, 264683, 264684, 264689, 18108359, 264691, 33657023, 264692, 35695763, 264629, 35695955, 264831, 264835, 264636, 264637, 56182323, 264639, 22729002, 264584
2139	94843882 (4277, 4278)	Novel Protein sim. GBank #J385082 [homoCA437135] - (Y18350) U2 snRNP auxiliary factor, large subunit [Nicotiana glauca]	UNCLASSIFIED	2646369, 264910, 265009, 264591, 264758, 52646317, 265011, 87188559, 264601, 18108351, 264446, 264683, 264684, 264689, 18108359, 264691, 33657023, 264692, 35695763, 264629, 35695955, 264831, 264835, 264636, 264637, 56182323, 264639, 22729002, 264584

21140	87645655 (4278, 4280)	Novel Protein sim. GBank gji4117293(gji4A020418)- (AC007019) unknown protein [Arabidopsis thaliana]		UNCLASSIFIED	264488, 264259, 29331824, 264104, 264105, 264106, 265006, 264793, 265018, 264448, 264501, 264963, 264975, 264977, 265021, 33657023, 27166533, 35686588, 264636, 264558, 264557, 264559, 264565
21141	79623986 (4281, 4282)	Novel Protein sim. GBank gji2135766(gjiS33382 - nucin SAC (dñe JER47) - human (fragment)		UNCLASSIFIED	263978
21142	60041222 (4283, 4284)	Novel Protein sim. GBank gji2078483 (L433200) - antifreeze glycoprotein AFGP polypeptide precursor [Boreogadus saius]		UNCLASSIFIED	264606, 60433356, 264686
21143	94140051 (4285, 4286)	Novel Protein sim. GBank gji1235766(gjiS33382 - nucin SAC (dñe JER47) - human (fragment)		UNCLASSIFIED	263978
21144	943200114 (4287, 4288)	Novel Protein sim. GBank gji2078483 (L433200) - antifreeze glycoprotein AFGP polypeptide precursor [Boreogadus saius]		UNCLASSIFIED	264606, 60433356, 264686
21145	20564305 (4289, 4290)	Novel Protein sim. GBank gji175571 (U53341) - short region of weak similarity to bovine membrane receptor p63 (P1R32653) [Ctenorhynchus elegans]		UNCLASSIFIED	264606, 264908, 264491, 265011, 87168559
21146	87010515 (4291, 4292)	Novel Protein sim. GBank gji3080398(jemb)CAA187181)- (AL027603) putative protein [Arabidopsis thaliana]		UNCLASSIFIED	264906, 265019, 264288, 264768, 18105765
21147	86432911 (4293, 4294)	Novel Protein sim. GBank gji2078483 (L433200) - antifreeze glycoprotein AFGP polypeptide precursor [Boreogadus saius]		UNCLASSIFIED	2196767, 5811576, 35686588, 65274791, 22279002
21148	60404811 (4295, 4296)	Novel Protein sim. GBank gji2078483 (L433200) - antifreeze glycoprotein AFGP polypeptide precursor [Boreogadus saius]		UNCLASSIFIED	263978
21149	87362022 (4297, 4298)	Novel Protein sim. GBank gji175571 (U53341) - short region of weak similarity to bovine membrane receptor p63 (P1R32653) [Ctenorhynchus elegans]		UNCLASSIFIED	264606, 264908, 264491, 265011, 87168559
21150	94140059 (4299, 4300)	Novel Protein sim. GBank gji2078483 (L433200) - antifreeze glycoprotein AFGP polypeptide precursor [Boreogadus saius]		UNCLASSIFIED	264606, 264908, 264491, 265011, 87168559
21151	95353241 (4301, 4302)	Novel Protein sim. GBank gji2078483 (L433200) - antifreeze glycoprotein AFGP polypeptide precursor [Boreogadus saius]		UNCLASSIFIED	264606, 264908, 264491, 265011, 87168559
21152	76321640 (4303, 4304)	Novel Protein sim. GBank gji3452473 (AF084205) - serine/threonine protein kinase TAO1 [Rattus norvegicus]		UNCLASSIFIED	80710394, 18108385, 356007, 264549, 265011, 18108351, 18108366, 18108374, 18108388

2153	88313371 (4305, 4306)	Novel Protein sim. GBank g1225150p[112952650] - g14758704[erfNP_004215.1]p[AS1 - MFI-amplified sequences with leucine-rich tandem repeats 1	Contains protein domain (PF00360) Leucine Rich Repeat	glycoprotein	264488, 265594, 5564842, 2271896; 2271898, 2271899, 264905, 264906, 3569053, 264906, 264909, 264905, 264905, 264907, 264909, 264909, 55182455, 264510, 264511, 264512, 264516, 87168474, 87168559, 265017, 265019, 264780, 264288, 264369, 264768, 264687, 264769, 52644229, 21606766, 21968768, 35699917, 33657023, 33657109, 33695855, 264631, 264632, 264635, 264635, 264639, 15108385, 264483, 264504, 264486
2154	87408034 (4307, 4308)	Novel Protein sim. GBank g1225150p[112952650] - chorion protein B11 [Bombyx mori]		UNCLASSIFIED	96994075, 264494, 265009, 265019, 264288, 21506767, 35695917
2155	87434072 (4309, 4310)			UNCLASSIFIED	18106392, 18106396, 22278986, 264259, 29331824, 265008, 265010, 265011, 265017, 265019, 264288, 264686, 265020, 264693, 264695, 35182253
2156	84265205 (4311, 4312)	Novel Protein sim. GBank g1307096[ACD04574] - like similar to AF026504 (PID:32555183) [Homo sapiens]			264697, 264698
2157	87316344 (4313, 4314)	Novel Protein sim. GBank g11076211[pi150755 - hypothetical protein VSP-3 - Chlamydomonas reinhardtii]		UNCLASSIFIED	264691
2158	86444218 (4315, 4316)	Novel Protein sim. GBank g1465084[5b]BAA77027.1] - (AB026190) Kelch motif containing protein [Homo sapiens]	Contains protein domain (PF00361) - dina_rna_bund		29331822, 264112, 265009, 264691, 33657023, 264634
2159	80083729 (4317, 4318)	Novel Protein sim. GBank g12879925[6]BAA424826] - (AB027897) KIAA0437 [Homo sapiens]			264634
2160	16205974 (4319, 4320)			UNCLASSIFIED	265008
2161	87759131 (4321, 4322)	Novel Protein sim. GBank g11504006[6]BAA13202] - (C88956) similitudo human ZFY protein [Homo sapiens]		UNCLASSIFIED	95374572, 264408, 264905, 264906, 264907, 264908, 52644445, 264909, 265007, 264910, 264951, 264952, 264953, 53812038, 264956, 264758, 265011, 264500, 264762, 264763, 264683, 264764, 264288, 264766, 264686, 264768, 264768, 264698, 265020, 264691, 264693, 264695, 265017, 265019, 264780, 264557, 264631, 264635, 15108385, 264563, 264556, 264597
2163	95417155 (4325, 4326)	Novel Protein sim. GBank g13076537[emb]CAA98270] - (Z73974) cDNA EST M-2916.3 comes from this gene; cDNA EST M-2916.5 comes from this gene [Caenorhabditis elegans]		UNCLASSIFIED	56182575, 22278956, 264083, 264683, 33657023, 65244620, 60432113
2164	80589456 (4327, 4328)		Contains protein domain (PF01006) Hepatitis C virus non-structural protein NS4a	collagen	264603, 264637, 264565

2165	(A)329169 (4329, 4330)	Novel Protein sim. GBank g11065794 (J41107) - No definition line found [Caenorhabditis elegans]	UNCLASSIFIED	264599, 26331822, 22278998, 22278997, 22278998, 264599, 26331822, 26331824, 26331825, 26331826, 26331827, 26331828, 264606, 26331830, 56182435, 265009, 21906794, 264369, 21906795, 264607, 264608, 264369, 21906796, 264609, 264610, 265020, 265021, 264691, 264692, 31857023, 65274620, 35695985, 264556, 60170904, 83373044, 60432113, 22278002, 264567
2166	(B)718934 (4331, 4332)	Novel Protein sim. GBank g1106572 (emb)CA71816] - (Y15895) ubiquitin activating enzyme [Drosophila melanogaster]	ubiquitin	26331828, 52644045, 265018, 265019, 264369, 21906765, 21906767, 21906768, 21906769, 265021, 265022, 264693, 27486262, 35695763, 19106376, 56526486, 87168518, 264567
2167	(B)716864 (4333, 4334)	Novel Protein sim. GBank g1122471 (g)g1AA20840] - (AB002384) Klu00306 [Homo sapiens]	UNCLASSIFIED	56182575, 35696286, 26331824, 26331825, 26146498, 56182435, 265009, 265009, 264592, 264593, 33657402, 33109954, 265011, 265017, 265019, 16108391, 264389, 21906704, 21906765, 21906788, 23148627, 21906789, 52654130, 33657169, 35696423, 16108391, 16108394, 19106385, 60432113, 264567
2168	(B)6999334 (4335, 4336)	Novel Protein sim. GBank g11432140 (g)g1AAD15748] - (AF047690) ATP-binding cassette protein M-ABC1 [Homo sapiens]	Contains protein domain (PF00684) - transport region.	66714117, 26331827, 264607, 264511, 264591, 265018, 584764, 264883, 264786, 264788, 264556
2169	(B)786937 (4337, 4338)	Novel Protein sim. GBank g1106521 (g)g1AAD39741.1] (AF10536) - (AF105365) X-CJ collagen transporter KCC4 [Homo sapiens]	UNCLASSIFIED	264620, 264555, 264559
2170	(B)4141033 (4339, 4340)	Novel Protein sim. GBank g1106521 (g)g1AAD39741.1] (AF10536) - (AF105365) X-CJ collagen transporter KCC4 [Homo sapiens]	UNCLASSIFIED	65274572, 36162575, 22278997, 22278998, 264259, 26331825, 264409, 264906, 56182435, 60433438, 55812038, 264596, 55811386, 265019, 264762, 254763, 264448, 264764, 264694, 264288, 264766, 264685, 56181562, 264688, 55811957, 265020, 264335, 264691, 33657109, 60431528, 16108374, 35696423, 55811576, 65274791, 264634, 264639, 264558, 67108516, 65274791, 264634, 264639, 264558, 67108516
2171	(B)0194050 (4341, 4342)	Novel Protein sim. GBank g11432140 (g)g1AAD15748] - (AF047690) ATP-binding cassette protein M-ABC1 [Homo sapiens]	UNCLASSIFIED	264369, 26331822, 26331824, 26331825, 26331826, 26331827, 26331828, 264606, 26331830, 56182435, 265009, 21906794, 264369, 21906795, 264607, 264608, 264369, 21906796, 264609, 264610, 265020, 265021, 264691, 264692, 31857023, 65274620, 35695985, 264556, 60170904, 83373044, 60432113, 22278002, 264567
2172	(B)5452460 (4343, 4344)	Novel Protein sim. GBank g11432140 (g)g1AAD15748] - (AF047690) ATP-binding cassette protein M-ABC1 [Homo sapiens]	UNCLASSIFIED	264369, 26331822, 26331824, 26331825, 26331826, 26331827, 26331828, 264606, 26331830, 56182435, 265009, 21906794, 264369, 21906795, 264607, 264608, 264369, 21906796, 264609, 264610, 265020, 265021, 264691, 264692, 31857023, 65274620, 35695985, 264556, 60170904, 83373044, 60432113, 22278002, 264567
2173	(B)7036740 (4345, 4346)	Novel Protein sim. GBank g11432140 (g)g1AAD15748] - (AF047690) ATP-binding cassette protein M-ABC1 [Homo sapiens]	UNCLASSIFIED	264369, 26331822, 26331824, 26331825, 26331826, 26331827, 26331828, 264606, 26331830, 56182435, 265009, 21906794, 264369, 21906795, 264607, 264608, 264369, 21906796, 264609, 264610, 265020, 265021, 264691, 264692, 31857023, 65274620, 35695985, 264556, 60170904, 83373044, 60432113, 22278002, 264567
2174	(B)5003288 (4347, 4348)	Novel Protein sim. GBank g12493778 (p)00945] Q35, CAEL - PUTATIVE [PUTATIVE COLLAGEN C9G5.5]	UNCLASSIFIED	264906, 35695985, 264555, 264557

2175	9432860 (3436, 4350)	Novel Protein sim. GBank g11253287 (U78455) - fibron-3 (Anous gadomus)		UNCLASSIFIED	264488, 3566826, 2028109, 20331826, 6043289, 3566802, 264109, 264508, 264509, 264505, 264506, 264507, 264508, 264509, 264510, 264511, 265006, 265007, 264512, 265008, 264910, 264581, 264594, 264595, 264598, 264758, 55812038, 265011, 264600, 264603, 204760, 204762, 264448, 264764, 264288, 264766, 264686, 264687, 21806768, 55811857, 35665917, 265020, 265022, 264691, 264692, 33857023, 264693, 264623, 264629, 55811576, 35669423, 65274791, 35669505, 264631, 264632, 264634, 264635, 264636, 264637, 264638, 264639, 264658, 18103393, 60432113, 264653, 264654, 264655, 264656, 26466, 26467, 52646942, 22778594, 35668286, 22778595, 25531829, 26331827, 35668592, 26331838, 33856870, 26331830, 264610, 33657402, 264758, 52644296, 87168559, 265016, 264689, 21508765, 21906767, 21906769, 35695917, 52644150, 264690, 33657023, 33657108, 52645129, 33657182, 27486261, 27486262, 33657348, 18108376, 18108377, 35695855, 87168518, 60432113, 264404, 22779000, 264485
2176	88223392 (4351, 4352)	Novel Protein sim. GBank g1724837 (p19194ALLUT, HUMAN - III ALU SUBFAMILY SQ WARNING ENTRY III)	Contains protein domain (PF00809)- Pentapeptide repeats (8 copies)	oncogene	264488, 3566826, 2028109, 20331826, 6043289, 3566802, 264109, 264508, 264509, 264510, 264511, 265006, 265007, 264512, 265008, 264910, 264581, 264594, 264595, 264598, 264758, 55812038, 265011, 264600, 264603, 204760, 204762, 264448, 264764, 264288, 264766, 264686, 264687, 21806768, 55811857, 35665917, 265020, 265022, 264691, 264692, 33857023, 264693, 264623, 264629, 55811576, 35669423, 65274791, 35669505, 264631, 264632, 264634, 264635, 264636, 264637, 264638, 264639, 264658, 18103393, 60432113, 264653, 264654, 264655, 264656, 26466, 26467, 52646942, 22778594, 35668286, 22778595, 25531829, 26331827, 35668592, 26331838, 33856870, 26331830, 264610, 33657402, 264758, 52644296, 87168559, 265016, 264689, 21508765, 21906767, 21906769, 35695917, 52644150, 264690, 33657023, 33657108, 52645129, 33657182, 27486261, 27486262, 33657348, 18108376, 18108377, 35695855, 87168518, 60432113, 264404, 22779000, 264485
2177	94128842 (4353, 4354)	Novel Protein sim. GBank g15454072 (p1NP_006416, 1pSLU7 - step II splicing factor SLU7)		kinase	18108392, 22278897, 22278595, 264095, 33657402, 265019, 264448, 264766, 264688, 21906767, 21906768, 21906769, 265021, 33857023, 18108370, 18108374, 60432113, 22779002
2178	87601557 (4355, 4356)	Novel Protein sim. GBank g1473407 (U08215) - NST-1 (Mus musculus)	Contains protein domain (P10012)- Hsp70 protein	eph	264488, 3566826, 2028109, 20331826, 6043289, 3566802, 264109, 264508, 264509, 264510, 264511, 265006, 265007, 264512, 265008, 264910, 264581, 264594, 264595, 264598, 264758, 55812038, 265011, 264600, 264603, 204760, 204762, 264448, 264764, 264288, 264766, 264686, 264687, 21806768, 55811857, 35665917, 265020, 265022, 264691, 264692, 33857023, 264693, 264623, 264629, 55811576, 35669423, 65274791, 35669505, 264631, 264632, 264634, 264635, 264636, 264637, 264638, 264639, 264658, 18103393, 60432113, 264653, 264654, 264655, 264656, 26466, 26467, 52646942, 22778594, 35668286, 22778595, 25531829, 26331827, 35668592, 26331838, 33856870, 26331830, 264610, 33657402, 264758, 52644296, 87168559, 265016, 264689, 21508765, 21906767, 21906769, 35695917, 52644150, 264690, 33657023, 33657108, 52645129, 33657182, 27486261, 27486262, 33657348, 18108376, 18108377, 35695855, 87168518, 60432113, 264404, 22779000, 264485
2179	87318275 (4357, 4358)			UNCLASSIFIED	264488, 3566826, 2028109, 20331826, 6043289, 3566802, 264109, 264508, 264509, 264510, 264511, 265006, 265007, 264512, 265008, 264910, 264581, 264594, 264595, 264598, 264758, 55812038, 265011, 264600, 264603, 204760, 204762, 264448, 264764, 264288, 264766, 264686, 264687, 21806768, 55811857, 35665917, 265020, 265022, 264691, 264692, 33857023, 264693, 264623, 264629, 55811576, 35669423, 65274791, 35669505, 264631, 264632, 264634, 264635, 264636, 264637, 264638, 264639, 264658, 18103393, 60432113, 264653, 264654, 264655, 264656, 26466, 26467, 52646942, 22778594, 35668286, 22778595, 25531829, 26331827, 35668592, 26331838, 33856870, 26331830, 264610, 33657402, 264758, 52644296, 87168559, 265016, 264689, 21508765, 21906767, 21906769, 35695917, 52644150, 264690, 33657023, 33657108, 52645129, 33657182, 27486261, 27486262, 33657348, 18108376, 18108377, 35695855, 87168518, 60432113, 264404, 22779000, 264485

2180	95351397 (4359, 4360)	Novel Protein sim. GBank g 31223 Thrip0494 KHB_DICD1 - MYOSIN HEAVY CHAIN KINASE B (MCKC B)	Contains protein domain (PF00400) WD domain, C-beta repeat	kinase	528444507, 22278934, 35096286, 22278997, 22278998, 284259, 28445000, 29331822, 29331823, 29331824, 29331825, 29331826, 29331827, 33855970, 284508, 284509, 284510, 284511, 29331830, 284906, 284910, 284905, 285007, 33857402, 55812038, 21906754, 87168474, 87168559, 285017, 285018, 285019, 284763, 284682, 284683, 284684, 284288, 284686, 21906765, 21906768, 21906769, 285020, 285021, 285022, 52644150, 33857023, 33857109, 27486285, 33857349, 18108374, 35695423, 35695555, 285981, 60170394, 18108385, 59526486, 87168518, 00432113, 22279000, 284482, 284566, 284557, 284486, 29331827, 284365, 18108376, 285564
2181	85164930 (4361, 4362)	Novel Protein sim. GBank g 510695 g 1544172D - HUMAN - TRANSCRIPTION FACTOR (TFIID 130 NO SUBUNIT (TAFL100))		kinase	22278998, 22278997, 22278999, 284349, 29331822, 58162435, 284112, 284764, 284288, 21906767, 21906768, 21906769, 33857109, 18108376, 60170394, 22279000, 22279002
2182	87637731 (4363, 4364)	Novel Protein sim. GBank g 542038 temb CAB4667811 - (AJ2434459) proteoglycophagyan (Leishmania major)		UNCLASSIFIED	22278998, 22278997, 22278999, 284349, 29331822, 58162435, 284112, 284764, 284288, 21906767, 21906768, 21906769, 33857109, 18108376, 60170394, 22279000, 22279002
2183	85460649 (4365, 4366)	Novel Protein sim. GBank g 3873408 g AAC7748211 - (U17129) unknown [Rhodococcus erythropolis]		Inf	29331822, 29331825, 29331826, 36182435, 285011, 284685, 284686, 21906768, 18108370, 2844529, 284631, 284636, 284557, 29331824, 284907, 66712502, 284757, 285019, 284288, 284692, 56529486
2184	87760650 (4367, 4368)	Novel Protein sim. GBank g 3114713 (AF061346) - Edp1 protein [Mus musculus]		Inf	29331822, 29331825, 29331826, 36182435, 285011, 284685, 284686, 21906768, 18108370, 2844529, 284631, 284636, 284557, 29331824, 284907, 66712502, 284757, 285019, 284288, 284692, 56529486
2185	87825483 (4369, 4370)	Novel Protein sim. GBank g 510695 g AD396061AF11381 - (AF113815) Prl17H2 domain-containing protein FHO5 (Homo sapiens)		ATPase-associated	264259, 29331822, 29331824, 29331826, 5618435, 284592, 55812038, 284760, 284786, 55811857, 33857023, 55811576, 33857023, 284508, 284509, 284510, 33859052, 284905, 284906, 284907, 284908, 284510, 284511, 285008, 284910, 284758, 285019, 284762, 284681, 284766, 284769, 35695917, 284692, 35696423, 284631, 284635, 284637, 18108388, 284586, 284486
2186	87738227 (4371, 4372)	Novel Protein sim. GBank g 2844625 temb CAB16972 - (AJ201811) putative protein [A. nidulans Italianus]		UNCLASSIFIED	264259, 29331822, 29331824, 29331826, 5618435, 284592, 55812038, 284760, 284786, 55811857, 33857023, 55811576, 33857023, 284508, 284509, 284510, 33859052, 284905, 284906, 284907, 284908, 284510, 284511, 285008, 284910, 284758, 285019, 284762, 284681, 284766, 284769, 35695917, 284692, 35696423, 284631, 284635, 284637, 18108388, 284586, 284486
2187	87388173 (4373, 4374)			UNCLASSIFIED	264259, 29331822, 29331824, 29331826, 5618435, 284592, 55812038, 284760, 284786, 55811857, 33857023, 55811576, 33857023, 284508, 284509, 284510, 33859052, 284905, 284906, 284907, 284908, 284510, 284511, 285008, 284910, 284758, 285019, 284762, 284681, 284766, 284769, 35695917, 284692, 35696423, 284631, 284635, 284637, 18108388, 284586, 284486
2188	87771708 (4375, 4376)	Novel Protein sim. GBank g 5107616 g AD401251AF14841 - (AF148413) contains similarity to histone deacetylases; Pfam PF00850, Score=13.3, E=5e-10, N=1 (Arabidopsis thaliana)		histone	18108388, 59594076, 284259, 29331824, 29331825, 66714117, 29331827, 284908, 29331830, 285018, 285020, 285021, 56182323, 284595, 22279000, 22279002
2189	85663573 (4377, 4378)	Novel Protein sim. GBank g 3452357 (AF075724) - Unknown [Legionella pneumophila]	Contains protein domain (PF01598) - O-methyltransferase		22278998, 284259, 29331826, 21906754, 284369, 284288, 283987

2190	87659197 (4374, 4380)	Novel Protein sm. GBank g 327599 P23015FRIAL_FAT - RIBONUCLEASE INHIBITOR		nucleoside	22278696, 22278599, 29331822, 29331824, 29331826, 265008, 264910, 60170831, 55812038, 5264298, 265010, 265018, 264885, 264688, 56181562, 21906769, 35685917, 265022, 60170394, 22279000
2191	95188928 (4381, 4382)	Novel Protein sm. GBank g 5273002 emb CA1846272.1 - (Y18603) XAP-5 like protein [Homo sapiens]			29331825, 29331826, 29331830, 264510, 264511, 264910, 264583, 264594, 284556, 264559
2192	11126316 (4383, 4384)	Novel Protein sm. GBank g 452600 sp P34400 ML10. CAEL - MIG-10 PROTEIN		Contains protein domain (PF00109) - PPI domain	264558
2193	94140073 (4385, 4386)	Novel Protein sm. GBank g 542038 emb CA1846680.1 - (AJ243460) ptaeophosphoglycan [Leishmania major]		UNCLASSIFIED	58151686, 29331825, 29331827, 224508, 264906, 265008, 264582, 60432225, 264288, 604684, 264768, 35695917, 3857023, 60431802, 60431822, 5810164, 55811576, 65371271, 35695955, 60431650, 56182321, 60432113, 264582
2194	21418714 (4387, 4388)	Novel Protein sm. GBank g 2773341 (AF040654) - putative protein phosphatase 1 nuclear targeting subunit [Rattus norvegicus]		UNCLASSIFIED	22278986, 22278698, 35696052, 265006, 21906754, 265017, 35695917, 265021, 265022, 35695955
2195	86083023 (4389, 4390)	Novel Protein sm. GBank g 2832763 emb CA11585.1 - (AJ009191) [prediction:method : [prediction:method : [match=desc.: [match=desc.: [match=desc.: [Drosophila melanogaster]		collagen	56182575, 35696286, 22278987, 22278989, 264259, 29331822, 60714117, 60432289, 29331827, 35696052, 29331828, 264598, 52644045, 56182435, 264510, 265007, 265008, 265009, 60433438, 55812038, 265010, 265011, 264448, 264288, 264686, 264687, 52644229, 21906785, 21906786, 21906787, 35695917, 265022, 264691, 33857023, 264693, 18108370, 18108376, 35696423, 55811576, 65274791, 35695955, 264636, 56182322, 18108385
2197	95073813 (4393, 4394)	Novel Protein sm. GBank g 4929567 sp AD34044.1 AF15180 - (AF151807) CGH49 protein [Homo sapiens]			21906787, 29148627, 55811937, 356959286, 265020, 22278998, 265021, 264259, 35697023, 264693, 29331824, 35696002, 29331825, 18108370, 35695955, 264113, 265009, 18108370, 60432228, 56182323, 33857402, 264693, 18108370, 21906794, 265018, 265019, 22278987, 264482, 264448, 264585, 264288, 264369
2198	88060691 (4395, 4396)	Novel Protein sm. GBank g 3548787 (AC005622) - R30953 [Homo sapiens]		UNCLASSIFIED	

2199	86054355 (4397, 4398)	Novel Protein sim. GBank g127939372 (A0202505) - hypothetical protein [Arabidopsis thaliana]			284105, 284110, 284112, 284688, 55811957, 33657023, 284992, 283987, 20281071, 36224846
2200	87455385 (4399, 4400)	Novel Protein sim. GBank g13043834(dj)g1A435461] - (AB011127) KIAA0555 sim [Homo sapiens]	struct		29331824, 284763, 284768
2201	84316872 (4401, 4402)	Novel Protein sim. GBank g139134706p1057314P4BXB ANKRA - PUTATIVE STEROID DEHYDROGENASE SPIN2	Contains protein domain (PF00108) short chain dehydrogenase	dehydrogenase	29331824, 28469052, 264905, 264907, 33657402, 5681356, 265017, 265018, 265019, 264288, 21968768, 35695917, 265020, 265022, 33657023, 33657109, 27486261, 18108370, 35696423, 35695655, 264555, 264556, 83373044, 87168518, 60432113
2202	91872385 (4403, 4404)	Novel Protein sim. GBank g15826665(emb)CA845767.1] - (AL000186) hypothetical protein [Homo sapiens]	UNCLASSIFIED	UNCLASSIFIED	254485, 264259, 29331824, 60432288, 35695952, 264905, 264909, 264592, 265017, 265018, 265019, 18108351, 264762, 264448, 264389, 264288, 264766, 21900765, 21900766, 264600, 264691, 264892, 33657109, 264634, 264636, 264535, 264639, 264538, 264539, 83373044, 18108385
2203	87761832 (4405, 4406)	Novel Protein sim. GBank g11728456p14629(RB25) RABIT - RAS-RELATED PROTEIN RAB-25	Contains protein domain (PF00071) Ras family	glycoprotein	27285426, 27285427, 264259, 29331822, 33643309, 35695917, 265018, 265019, 30331827, 26501828, 264910, 265010, 265011, 87168556, 265018, 265019, 264605, 264288, 21900769, 35695917, 33657023, 264892, 33657109, 35695763, 18108376, 264638, 22279000, 264586, 264587
2204	88088671 (4407, 4408)	Novel Protein sim. GBank g1121035(p1723348)CBT3_TAT GUANINE NUCLEOTIDE-BINDING PROTEIN G(T), ALPHA 3 SUBUNIT (GUSTOLUCIN ALPHA-3 CHAIN)	Contains protein domain (PF00503) - UNCLASSIFIED	UNCLASSIFIED	18108394, 18108397, 56182575, 60432049, 264259, 29331822, 29331824, 29331825, 29331826, 29331827, 264906, 265007, 265008, 265009, 80432278, 265010, 265011, 265018, 264683, 264285, 264380, 264686, 21900766, 21900768, 21900769, 35695917, 35695918, 35695919, 35695920, 65271791, 264634, 18108381, 18108384, 60432113, 22279002, 264683, 264586, 264681
2205	94147589 (4409, 4410)	Novel Protein sim. GBank g14589480(dj)3AA76768.1] - (AB023141) KAA0024 protein [Homo sapiens]	Contains protein domain (PF00086) Zinc finger, C2H2 type	dna_rna_bind	25331822, 56182181, 29331827, 35696052, 56844045, 265006, 265019, 56181582, 55811957, 265021, 33657023, 35695763, 35695855, 60170394, 60432113, 264586, 264906, 265009, 264691, 21900769, 264112, 265009, 264691, 18108385
2206	20622008 (4411, 4412)	Novel Protein sim. GBank g14557753(ne)NP_000372 [IPM1D1 - midline 1 protein]	Contains protein domain (PF00622) - SPRY domain	UNCLASSIFIED	264681
2207	87761870 (4413, 4414)	Novel Protein sim. GBank g13956746 (AF105228) - Iuflin [Bos taurus]	struct		18108374, 264634, 20281169

2220	95354165 (4433, 4440)	Novel Protein sim. GBank g10549154 (AC005625) - g10549154 [Homo sapiens]			264488, 18108394, 18108395, 35895286, 264259, 264907, 6043289, 264509, 264905, 264906, 264907, 5933130, 264908, 264909, 264510, 264511, 850007, 264512, 264910, 265093, 264583, 264594, 6043385, 264595, 264596, 264597, 264598, 264599, 264600, 264601, 264602, 264603, 264604, 264605, 264606, 264607, 264608, 264609, 264610, 264611, 264612, 264613, 264614, 264615, 264616, 264617, 264618, 264619, 264620, 264621, 264622, 264623, 264624, 264625, 264626, 264627, 264628, 264629, 264630, 264631, 264632, 264633, 264634, 264635, 264636, 264637, 264638, 18108385, 264483, 264566, 264486, 264567
2221	88060927 (4441, 4442)	Novel Protein sim. GBank g10549154 (AC005625) - R27328.1 [Homo sapiens]			
2222	84435892 (4443, 4444)			UNCLASSIFIED	264508, 265020, 35895655
2223	95391649 (4445, 4446)			UNCLASSIFIED	265010, 264685, 264690, 264693, 264628, 263974, 263976, 55911576, 264555, 264638, 83373044, 264483
2224	87398015 (4447, 4448)	Novel Protein sim. GBank g107600919 (C15719) cDNA EST EMBL C1583 comes from this gene. cDNA EST EMBL C1578 comes from this gene. cDNA EST y224a7.3 comes from this gene. cDNA EST y224a7.5 comes from this gene. cDNA ES...	Contains protein domain (PF01958) Domain of unknown function	UNCLASSIFIED	264259, 264509, 36162435, 265006, 265008, 265009, 264737, 21960754, 18106351, 264693, 18106374, 18106365
2225	85716484 (4449, 4450)	Novel Protein sim. GBank g11255647 (U53338) - C03E11.1 gene product [Caenorhabditis elegans]	Transport	Transport	22278994, 22278995, 22278999, 33944043, 264600, 265019, 21905765, 21905769
2226	86978953 (4451, 4452)	Novel Protein sim. GBank g10498948 (AL049894) hypothetical protein [Homo sapiens]			264259, 26531822, 29331824, 29331825, 29331827, 264508, 264606, 265007, 264691, 264634, 264486
2227	87721135 (4453, 4454)			UNCLASSIFIED	22278999, 265006, 265008, 18108354, 29148629, 29148794, 27486261, 18106374, 264637, 18106394
2228	91227337 (4455, 4456)	Novel Protein sim. GBank g1068076 (U16800) - ribonucleoprotein [Panopus tawui]	Contains protein domain (PF06036) RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain)	diva_ma_bind	264637, 18106394, 29331825, 68714117, 264693, 263972, 264633, 63373044, 264563
2229	88060931 (4457, 4458)	Novel Protein sim. GBank g10549155 (AC005625) - R27328.2 [Homo sapiens]		UNCLASSIFIED	

2238	94098857 (4475, 4476)		Contains protein domain (PF00286) - Viral coat protein	264508, 264907, 264629, 264634, 264644
2239	87196868 (4477, 4478)			20331825, 264009, 264008, 33657109, 18103870, 18103874, 264547, 264559, 264485, 55274572, 56192725, 33686286, 22278997, 22278998, 264259, 20331827, 336956032, 264506, 52944045, 56182435, 264511, 265007, 265008, 265009, 60433366, 60433438, 55812038, 21906754, 33657084, 55811386, 265016, 265019, 18108391, 264683, 264288, 264766, 264687, 264688, 264769, 21906765, 21906768, 21906769, 33695917, 265021, 265022, 60170615, 52644150, 33657023, 33657182, 33657340, 33695763, 18106370, 33696423, 33695855, 26455918, 22278990
2240	94121471 (4479, 4480)		Novel Protein sim. GBank gij0992311 (AF051240) - probable ubiquitin-conjugating enzyme E2 [Picea mariana]	UNCLASSIFIED
2241	80091951 (4481, 4482)		Novel Protein sim. GBank	22278995, 22278996, 22278997, 22278998, 204259, 20331822, 20331824, 20331826, 20331827, 20331828, 21906759, 265007, 285009, 264566, 21906754, 265010, 265011, 265017, 265018, 265019, 264446, 264389, 264288, 52644229, 21906765, 21906766, 21906767, 21906768, 21906769, 265020, 265021, 33657105, 27486262, 27486264, 18108374, 33695855, 264634, 264637, 56182323, 83373044, 56526468, 87168618, 264564
2242	91228075 (4483, 4484)		Novel Protein sim. GBank gij2494312ipP70541IE2BG_RAT - TRANSLATION INITIATION FACTOR EIF-2B GAMMA SUBUNIT (EIF-2B GDP-GTP EXCHANGE FACTOR)	UNCLASSIFIED synthase
2243	78902026 (4485, 4486)		Novel Protein sim. GBank gij2251143 (AF018417) - Similar to B2IP transcription factor [Caenorhabditis elegans]	UNCLASSIFIED
2244	65723227 (4487, 4488)		Novel Protein sim. GBank gij410340 (U00043) - similar to beta-mannosyltransferase [Caenorhabditis elegans]	UNCLASSIFIED
2245	65318545 (4489, 4490)		Contains protein domain (PF00534) - Glycosyl transferases group 1	13654158, 22278995, 22278996, 22278997, 22278998, 25319422, 20331824, 20331827, 264907, 264511, 264603, 264519, 21906754, 265011, 204603, 264548, 264687, 21906765, 21906768, 21906769, 55811397, 265022, 264681, 264690, 33696423, 264638, 18108387, 60432113, 22279000, 22279002, 264566

2246	94848710 (4481, 4482)	Novel Protein sim. GBank g1496506[emb]BA37336.1] - (A8026069) activator of S phase kinase [Homo sapiens]	Contains protein domain (PF00153) - Mitochondrial cancer proteins	transport	55274572, 22278995, 35692665, 22278996, 22278997, 22278998, 22278999, 269229, 35696052, 269008, 60433336, 33109854, 67168559, 265018, 265019, 264298, 21906765, 21906767, 21906768, 21906769, 55811957, 35695917, 255020, 255022, 27466294, 18108370, 18108374, 65274791, 35695955, 60432113
2247	87862542 (4493, 4494)	Novel Protein sim. GBank g165406[emb]CAA55337] - (Q83413) U88 [Human herpesvirus 6]	UNCLASSIFIED		52945156, 52646365, 52645080, 35696052, 33656970, 52649317, 33657084, 35697109, 21906768, 21906769, 35695917, 33657109, 52645129, 33657182, 27466261, 27466282, 33657349, 27466255, 18108387
2248	95412986 (4495, 4496)	Novel Protein sim. GBank g175502[ref]NP_004123.1[pr]IAP - hyaluronan-binding protein 2	Contains protein domain (PF00069) - Trypsin	cathepsin	264488, 264259, 264407, 26331830, 264909, 265007, 265009, 264596, 21906754, 65274444, 264803, 265019, 264732, 264448, 264288, 264689, 21906706, 55811957, 265021, 264491, 18108374, 264634, 264635, 264636, 264637, 264638, 264639, 264640, 264641, 264642, 264643, 264644, 264645, 264646, 264647, 264648, 264649, 264650, 264651, 264652, 264653, 264654, 264655, 264656, 264657, 264658, 264659, 18108383, 83373044, 18108386, 264485
2249	94656562 (4487, 4488)	Novel Protein sim. GBank gH4035461 (AF107722) - TCST11 [Trypanosoma cruzi]	Contains protein domain (PF00515) - TPR Domain	eph	264756, 264626, 264536, 264637
2250	78927508 (4499, 4500)	Novel Protein sim. GBank g1328140[emb]CAA21241] - (AL031882) valyl-tRNA synthetase, mitochondrial precursor [Schistosoma haematophyllum]		UNCLASSIFIED	264908, 18108374
2251	87365853 (4501, 4502)	Novel Protein sim. GBank g13218467[emb]CAA07090.1] - (AJ005529) putative phosphatase [Gallus gallus]		UNCLASSIFIED	264229, 35696052, 264508, 56182435, 265009, 264592, 254593, 264790, 264448, 264684, 264283, 264690, 264628, 55811576, 264555, 264556, 264557, 264559, 264559, 264560
2252	87735887 (4503, 4504)	Novel Protein sim. GBank g14929325[emb]AA333955.1[AF14531] vacuolar ATP synthase subunit D [Homo sapiens]	Contains protein domain (PF01813) - ATP synthase subunit D	synthase	264032, 264094, 264459, 26331822, 66714117, 25331628, 264102, 261003, 264074, 264105, 264106, 264112, 264511, 264512, 264513, 264514, 264515, 264516, 21906757, 21906768, 264691, 263974, 263975, 264486, 264487, 264488, 264489, 264490, 264491, 264492, 264493, 264494, 264495, 264496, 264497, 264498, 264499, 264500, 264501, 264502, 264503, 264504, 264505, 264506, 264507, 264508, 264509, 264510, 264511, 264512, 264513, 264514, 264515, 264516, 264517, 264518, 264519, 264520, 264521, 264522, 264523, 264524, 264525, 264526, 264527, 264528, 264529, 264530, 264531, 264532, 264533, 264534, 264535, 264536, 264537, 264538, 264539, 264540, 264541, 264542, 264543, 264544, 264545, 264546, 264547, 264548, 264549, 264550, 264551, 264552, 264553, 264554, 264555, 264556, 264557, 264558, 264559, 264560, 264561, 264562, 264563, 264564, 264565, 264566, 264567, 264568, 264569, 264570, 264571, 264572, 264573, 264574, 264575, 264576, 264577, 264578, 264579, 264580, 264581, 264582, 264583, 264584, 264585, 264586, 264587, 264588, 264589, 264590, 264591, 264592, 264593, 264594, 264595, 264596, 264597, 264598, 264599, 264600, 264601, 264602, 264603, 264604, 264605, 264606, 264607, 264608, 264609, 264610, 264611, 264612, 264613, 264614, 264615, 264616, 264617, 264618, 264619, 264620, 264621, 264622, 264623, 264624, 264625, 264626, 264627, 264628, 264629, 264630, 264631, 264632, 264633, 264634, 264635, 264636, 264637, 264638, 264639, 264640, 264641, 264642, 264643, 264644, 264645, 264646, 264647, 264648, 264649, 264650, 264651, 264652, 264653, 264654, 264655, 264656, 264657, 264658, 264659, 264660, 264661, 264662, 264663, 264664, 264665, 264666, 264667, 264668, 264669, 264670, 264671, 264672, 264673, 264674, 264675, 264676, 264677, 264678, 264679, 264680, 264681, 264682, 264683, 264684, 264685, 264686, 264687, 264688, 264689, 264690, 264691, 264692, 264693, 264694, 264695, 264696, 264697, 264698, 264699, 264700, 264701, 264702, 264703, 264704, 264705, 264706, 264707, 264708, 264709, 264710, 264711, 264712, 264713, 264714, 264715, 264716, 264717, 264718, 264719, 264720, 264721, 264722, 264723, 264724, 264725, 264726, 264727, 264728, 264729, 264730, 264731, 264732, 264733, 264734, 264735, 264736, 264737, 264738, 264739, 264740, 264741, 264742, 264743, 264744, 264745, 264746, 264747, 264748, 264749, 264750, 264751, 264752, 264753, 264754, 264755, 264756, 264757, 264758, 264759, 264760, 264761, 264762, 264763, 264764, 264765, 264766, 264767, 264768, 264769, 264770, 264771, 264772, 264773, 264774, 264775, 264776, 264777, 264778, 264779, 264780, 264781, 264782, 264783, 264784, 264785, 264786, 264787, 264788, 264789, 264790, 264791, 264792, 264793, 264794, 264795, 264796, 264797, 264798, 264799, 264800, 264801, 264802, 264803, 264804, 264805, 264806, 264807, 264808, 264809, 264810, 264811, 264812, 264813, 264814, 264815, 264816, 264817, 264818, 264819, 264820, 264821, 264822, 264823, 264824, 264825, 264826, 264827, 264828, 264829, 264830, 264831, 264832, 264833, 264834, 264835, 264836, 264837, 264838, 264839, 264840, 264841, 264842, 264843, 264844, 264845, 264846, 264847, 264848, 264849, 264850, 264851, 264852, 264853, 264854, 264855, 264856, 264857, 264858, 264859, 264860, 264861, 264862, 264863, 264864, 264865, 264866, 264867, 264868, 264869, 264870, 264871, 264872, 264873, 264874, 264875, 264876, 264877, 264878, 264879, 264880, 264881, 264882, 264883, 264884, 264885, 264886, 264887, 264888, 264889, 264890, 264891, 264892, 264893, 264894, 264895, 264896, 264897, 264898, 264899, 264900, 264901, 264902, 264903, 264904, 264905, 264906, 264907, 264908, 264909, 264910, 264911, 264912, 264913, 264914, 264915, 264916, 264917, 264918, 264919, 264920, 264921, 264922, 264923, 264924, 264925, 264926, 264927, 264928, 264929, 264930, 264931, 264932, 264933, 264934, 264935, 264936, 264937, 264938, 264939, 264940, 264941, 264942, 264943, 264944, 264945, 264946, 264947, 264948, 264949, 264950, 264951, 264952, 264953, 264954, 264955, 264956, 264957, 264958, 264959, 264960, 264961, 264962, 264963, 264964, 264965, 264966, 264967, 264968, 264969, 264970, 264971, 264972, 264973, 264974, 264975, 264976, 264977, 264978, 264979, 264980, 264981, 264982, 264983, 264984, 264985, 264986, 264987, 264988, 264989, 264990, 264991, 264992, 264993, 264994, 264995, 264996, 264997, 264998, 264999, 265000, 265001, 265002, 265003, 265004, 265005, 265006, 265007, 265008, 265009, 265010, 265011, 265012, 265013, 265014, 265015, 265016, 265017, 265018, 265019, 265020, 265021, 265022, 265023, 265024, 265025, 265026, 265027, 265028, 265029, 265030, 265031, 265032, 265033, 265034, 265035, 265036, 265037, 265038, 265039, 265040, 265041, 265042, 265043, 265044, 265045, 265046, 265047, 265048, 265049, 265050, 265051, 265052, 265053, 265054, 265055, 265056, 265057, 265058, 265059, 265060, 265061, 265062, 265063, 265064, 265065, 265066, 265067, 265068, 265069, 265070, 265071, 265072, 265073, 265074, 265075, 265076, 265077, 265078, 265079, 265080, 265081, 265082, 265083, 265084, 265085, 265086, 265087, 265088, 265089, 265090, 265091, 265092, 265093, 265094, 265095, 265096, 265097, 265098, 265099, 265100, 265101, 265102, 265103, 265104, 265105, 265106, 265107, 265108, 265109, 265110, 265111, 265112, 265113, 265114, 265115, 265116, 265117, 265118, 265119, 265120, 265121, 265122, 265123, 265124, 265125, 265126, 265127, 265128, 265129, 265130, 265131, 265132, 265133, 265134, 265135, 265136, 265137, 265138, 265139, 265140, 265141, 265142, 265143, 265144, 265145, 265146, 265147, 265148, 265149, 265150, 265151, 265152, 265153, 265154, 265155, 265156, 265157, 265158, 265159, 265160, 265161, 265162, 265163, 265164, 265165, 265166, 265167, 265168, 265169, 265170, 265171, 265172, 265173, 265174, 265175, 265176, 265177, 265178, 265179, 265180, 265181, 265182, 265183, 265184, 265185, 265186, 265187, 265188, 265189, 265190, 265191, 265192, 265193, 265194, 265195, 265196, 265197, 265198, 265199, 265200, 265201, 265202, 265203, 265204, 265205, 265206, 265207, 265208, 265209, 265210, 265211, 265212, 265213, 265214, 265215, 265216, 265217, 265218, 265219, 265220, 265221, 265222, 265223, 265224, 265225, 265226, 265227, 265228, 265229, 265230, 265231, 265232, 265233, 265234, 265235, 265236, 265237, 265238, 265239, 265240, 265241, 265242, 265243, 265244, 265245, 265246, 265247, 265248, 265249, 265250, 265251, 265252, 265253, 265254, 265255, 265256, 265257, 265258, 265259, 265260, 265261, 265262, 265263, 265264, 265265, 265266, 265267, 265268, 265269, 265270, 265271, 265272, 265273, 265274, 265275, 265276, 265277, 265278, 265279, 265280, 265281, 265282, 265283, 265284, 265285, 265286, 265287, 265288, 265289, 265290, 265291, 265292, 265293, 265294, 265295, 265296, 265297, 265298, 265299, 265300, 265301, 265302, 265303, 265304, 265305, 265306, 265307, 265308, 265309, 265310, 265311, 265312, 265313, 265314, 265315, 265316, 265317, 265318, 265319, 265320, 265321, 265322, 265323, 265324, 265325, 265326, 265327, 265328, 265329, 265330, 265331, 265332, 265333, 265334, 265335, 265336, 265337, 265338, 265339, 265340, 265341, 265342, 265343, 265344, 265345, 265346, 265347, 265348, 265349, 265350, 265351, 265352, 265353, 265354, 265355, 265356, 265357, 265358, 265359, 265360, 265361, 265362, 265363, 265364, 265365, 265366, 265367, 265368, 265369, 265370, 265371, 265372, 265373, 265374, 265375, 265376, 265377, 265378, 265379, 265380, 265381, 265382, 265383, 265384, 265385, 265386, 265387, 265388, 265389, 265390, 265391, 265392, 265393, 265394, 265395, 265396, 265397, 265398, 265399, 265400, 265401, 265402, 265403, 265404, 265405, 265406, 265407, 265408, 265409, 265410, 265411, 265412, 265413, 265414, 265415, 265416, 265417, 265418, 265419, 265420, 265421, 265422, 265423, 265424, 265425, 265426, 265427, 265428, 265429, 265430, 265431, 265432, 265433, 265434, 265435, 265436, 265437, 265438, 265439, 265440, 265441, 265442, 265443, 265444, 265445, 265446, 265447, 265448, 265449, 265450, 265451, 265452, 265453, 265454, 265455, 265456, 265457, 265458, 265459, 265460, 265461, 265462, 265463, 265464, 265465, 265466, 265467, 265468, 265469, 265470, 265471, 265472, 265473, 265474, 265475, 265476, 265477, 265478, 265479, 265480, 265481, 265482, 265483, 265484, 265485, 265486, 265487, 265488, 265489, 265490, 265491, 265492, 265493, 265494, 265495, 265496, 265497, 265498, 265499, 265500, 265501, 265502, 265503, 265504, 265505, 265506, 265507, 265508, 265509, 265510, 265511, 265512, 265513, 265514, 265515, 265516, 265517, 265518, 265519, 265520, 265521, 265522, 265523, 265524, 265525, 265526, 265527, 265528, 265529, 265530, 265531, 265532, 265533, 265534, 265535, 265536, 265537, 265538, 265539, 265540, 265541, 265542, 265543, 265544, 265545, 265546, 265547, 265548, 265549, 265550, 265551, 265552, 265553, 265554, 265555, 265556, 265557, 265558, 265559, 265560, 265561, 265562, 265563, 265564, 265565, 265566, 265567, 265568, 265569, 265570, 265571, 265572, 265573, 265574, 265575, 265576, 265577, 265578, 265579, 265580, 265581, 265582, 265583, 265584, 265585, 265586, 265587, 265588, 265589, 265590, 265591, 265592, 265593, 265594, 265595, 265596, 265597, 265598, 265599, 265600, 265601, 265602, 265603, 265604, 265605, 265606, 265607, 265608, 265609, 265610, 265611, 265612, 265613, 265614, 265615, 265616, 265617, 265618, 265619, 265620, 265621, 265622, 265623, 265624, 265625, 265626, 265627, 265628, 265629, 265630, 265631, 265632, 265633, 265634, 265635, 265636, 265637, 265638, 265639, 265640, 265641, 265642, 265643, 265644, 265645, 265646, 265647, 265648, 265649, 265650, 265651, 265652, 265653, 265654, 265655, 265656, 265657, 265658, 265659, 265660, 265661, 265662, 265663, 265664, 265665, 265666, 265667, 265668, 265669, 265670, 265671, 265672, 265673, 265674, 265675, 265676, 265677, 265678, 265679, 265680, 265681, 265682, 265683, 265684, 265685, 265686, 265687, 265688, 265689, 265690, 265691, 265692, 265693, 265694, 265695, 265696, 265697, 265698, 265699, 265700, 265701, 265702, 265703, 265704, 265705, 265706, 265707, 265708, 265709, 265710, 265711, 265712, 265713, 265714, 265715, 265716, 265717, 265718, 265719, 265720, 265721, 265722, 265723, 265724, 265725, 265726, 265727, 265728, 265729, 265730, 265731, 265732, 265733, 265734, 265735, 265736, 265737, 265738, 265739, 265740, 265741, 265742, 265743, 265744, 265745, 265746, 265747, 265748, 265749, 265750, 265751, 265752, 26

2254	95320031 (4507, 4508) gl4502847[refNP_001271.1]pICRF - cold inducible RNA-binding protein	Novel Protein sim. GBank gl4502847[refNP_001271.1]pICRF - cold inducible RNA-binding protein	Contains protein domain (PF00076) - RNA recognition motif. (p.k.a. RRM, RBD, or RNP domain)	one_rna_bind	264569, 18108394, 18108398, 56162575, 56994075, 3566296, 22278999, 264094, 60432048, 264296, 26331822, 29331824, 29331825, 29331826, 60432286, 28331827, 28331828, 28331829, 28331830, 2844008, 2844009, 2844010, 2844011, 2844012, 2844013, 2844014, 2844015, 2844016, 66712602, 2845008, 264909, 264510, 265008, 264511, 265007, 265008, 265009, 60770831, 60432228, 60433358, 60433438, 264758, 95558542, 265010, 265011, 87168559, 265017, 265018, 265019, 264448, 264764, 264288, 284388, 264766, 264688, 264768, 264769, 21906765, 21906767, 55811957, 264691, 33857023, 264692, 18106362, 65274620, 263989, 264628, 18108370, 60431528, 263972, 264629, 18108372, 18108377, 18108379, 55811576, 35696423, 33099853, 264630, 264634, 264635, 264636, 264558, 263981, 264638, 56162323, 60770394, 264636, 18106381, 18106382, 60770404, 18106385, 87168516, 60432113, 264637, 264638, 264639, 264640, 264641, 264642, 264643, 264644, 264645, 264646, 264647, 264648, 264649, 18108391
2255	91010546 (4509, 4510) Novel Protein sim. GBank gl5541865[emb(CAB51072.1)] (AL096859) hypothetical protein [Homo sapiens]	Novel Protein sim. GBank gl5541865[emb(CAB51072.1)] (AL096859) hypothetical protein [Homo sapiens]	Contains protein domain (PF00076) - RNA recognition motif. (p.k.a. RRM, RBD, or RNP domain)	rrnd	65274572, 56162575, 22278997, 22278999, 264259, 29331822, 29331825, 29331828, 29331827, 29331828, 294508, 264505, 264906, 264907, 66712502, 264508, 56182435, 264510, 264511, 265008, 264593, 264595, 21906754, 33109954, 87168474, 265011, 265017, 265019, 264682, 264764, 264369, 264284, 264768, 264685, 264686, 264768, 21906765, 21906766, 21906768, 21906769, 265020, 60770615, 52544190, 264690, 264692, 264693, 33857109, 33857349, 264632, 264636, 32644332, 56162323, 22278999, 22279002, 264768, 264688, 18108374
2256	87020531 (4511, 4512) Novel Protein sim. GBank gl337174[dbJBJ-A3.1656] - (AB014580) KIAA0680 protein [Homo sapiens]	Novel Protein sim. GBank gl337174[dbJBJ-A3.1656] - (AB014580) KIAA0680 protein [Homo sapiens]		UNCLASSIFIED	
2257	80088235 (4513, 4514) Novel Protein sim. GBank gl3025446 (AC004528) - R32184_2 [Homo sapiens]	Novel Protein sim. GBank gl3025446 (AC004528) - R32184_2 [Homo sapiens]	Contains protein domain (PF00060) - Ligand-gated ion channel	UNCLASSIFIED	22278998, 22278999, 264681, 21906765, 21906766, 264467
2258	86090516 (4515, 4516) Novel Protein sim. GBank gl3025446 (AC004528) - R32184_2 [Homo sapiens]	Novel Protein sim. GBank gl3025446 (AC004528) - R32184_2 [Homo sapiens]	Contains protein domain (PF00060) - Ligand-gated ion channel	UNCLASSIFIED	264608, 264592, 264764

2255	95364155 (4517, 4518)	Novel Protein sim. GBank g14384140[emb]CAB43278.11 - (AL050110) hypothetical protein [Homo sapiens]	UNCLASSIFIED	18103096, 65274572, 56183515, 22278997, 22278998, 384295, 26331822, 26331823, 264505, 69712502, 264508, 264509, 56183435, 265007, 265008, 60432279, 56175094, 57186559, 18108351, 284448, 264683, 284288, 264358, 56181562, 265021, 60170615, 264690, 33657109, 60431528, 18108374, 52644332, 56182323, 18108395, 22278900, 22278902, 264482
2260	88064119 (4519, 4520)	Novel Protein sim. GBank g31080665 (AC0304514) - similar to 14-podin proteins AB008086 (P1D3J252225) [Homo sapiens]	-oxidase	56182575, 255020, 264605, 264606, 264608, 35964423, 264511, 264635, 55812038, 264758, 265018, 265019, 264605, 264760, 264583
2261	88074157 (4521, 4522)	Novel Protein sim. GBank g31333455[emb]CAA161381 - (AL021306) predicted protein [F. GENENT] [Homo sapiens]	UNCLASSIFIED	56182575, 22278999, 26331822, 26331823, 264505, 69712502, 264508, 264509, 56175094, 57186559, 18108351, 284448, 264683, 284288, 264358, 56181562, 265021, 60170615, 264690, 33657109, 60431528, 18108374, 52644332, 56182323, 18108395, 22278900, 22278902, 264482
2262	91633232 (4523, 4524)	Novel Protein sim. GBank g1487759[gbAA031421.1AF12444 - (AF124440) MAGE tumor antigen O1 [Homo sapiens]	UNCLASSIFIED	56182575, 22278999, 26331822, 26331823, 264505, 69712502, 264508, 264509, 56175094, 57186559, 18108351, 284448, 264683, 284288, 264358, 56181562, 265021, 60170615, 264690, 33657109, 60431528, 18108374, 52644332, 56182323, 18108395, 22278900, 22278902, 264482
2263	87602495 (4525, 4526)	Novel Protein sim. GBank g31341897 (AC008572) - hypothetical protein [Rhabdopsis thaliana]	UNCLASSIFIED	56182575, 22278999, 26331822, 26331823, 264505, 69712502, 264508, 264509, 56175094, 57186559, 18108351, 284448, 264683, 284288, 264358, 56181562, 265021, 60170615, 264690, 33657109, 60431528, 18108374, 52644332, 56182323, 18108395, 22278900, 22278902, 264482
2264	87756525 (4527, 4528)	Novel Protein sim. GBank g11657801 (U96220) - unknown [Harmocystis exedens]	UNCLASSIFIED	56182575, 22278999, 26331822, 26331823, 264505, 69712502, 264508, 264509, 56175094, 57186559, 18108351, 284448, 264683, 284288, 264358, 56181562, 265021, 60170615, 264690, 33657109, 60431528, 18108374, 52644332, 56182323, 18108395, 22278900, 22278902, 264482
2265	86918653 (4529, 4530)	Novel Protein sim. GBank g1477072[pil]A48018 - mucin 7 precursor, salivary - human	UNCLASSIFIED	56182575, 22278999, 26331822, 26331823, 264505, 69712502, 264508, 264509, 56175094, 57186559, 18108351, 284448, 264683, 284288, 264358, 56181562, 265021, 60170615, 264690, 33657109, 60431528, 18108374, 52644332, 56182323, 18108395, 22278900, 22278902, 264482
2266	87773458 (4531, 4532)	Novel Protein sim. GBank g313150419 (AF062121) - partial CDS [Caenorhabditis elegans]	UNCLASSIFIED	56182575, 22278999, 26331822, 26331823, 264505, 69712502, 264508, 264509, 56175094, 57186559, 18108351, 284448, 264683, 284288, 264358, 56181562, 265021, 60170615, 264690, 33657109, 60431528, 18108374, 52644332, 56182323, 18108395, 22278900, 22278902, 264482

2207	87394538 (4533, 4534)	Novel Protein sim. GBank gji1560229(emt)CA20697.11- (AL031530) hypofunctional protein [Gonosaccharinomyces pombae]		UNCLASSIFIED	35998286, 264250, 29331824, 29331825, 35998502, 29331828, 264905, 264509, 264907, 264908, 264909, 264512, 265009, 264910, 264911, 264912, 265010, 265018, 264919, 264920, 264921, 264922, 264923, 264422, 264448, 264458, 264459, 264460, 52644228, 59595917, 264691, 35957023, 11083351, 264692, 264693, 264694, 264695, 11083362, 35657109, 35984523, 264634, 11083381, 87168518, 264588 264448, 264250, 264500, 264595, 265010, 265017, 264766, 18106385, 264466
2268	85693867 (4535, 4536)	Novel Protein sim. GBank glt28532p939169ALUZ_HUMAN - III ALU SUBFAMILY SB WARNING ENTRY III	cadherin	UNCLASSIFIED	56182575, 60432046, 265007, 265009, 264591, 87168558, 264605, 18108351, 21966764, 265020, 264629, 60431528, 264636, 18106395, 18108387, 60432113 264763
2270	86410327 (4539, 4540)				264909, 56182435, 265006, 56812038, 56811957, 35657023, 264893, 35657109, 56810764, 56811576, 56182323
2271	91010382 (4541, 4542)				53451163, 22276908, 22276909, 60432049, 264250, 29331822, 29331824, 29331825, 29331826, 29331827, 35998502, 264909, 265006, 264593, 60433438, 21906754, 265018, 264688, 21906765, 21906766, 21906767, 21906769, 265021, 265022, 60170615, 264691, 35657023, 264693, 35657109, 27489264, 18108376, 35696423, 35695955, 264630, 52644332, 264558, 56182323, 22276902
2272	84208220 (4543, 4544)	Novel Protein sim. GBank gji146176379 (AC005038) - similar to calcium-independent phospholipase A2, similar to AC004392 (PI0 g3367519) [Homo sapiens]	Contains protein domain (PF00462) Glutaredoxin	UNCLASSIFIED	52645156, 22276987, 22276989, 52645080, 29331824, 29331825, 29331826, 29331827, 29331828, 264905, 264908, 52644045, 264511, 265008, 265009, 80170831, 264591, 21906754, 33105954, 265011, 265018, 11083351, 264446, 264288, 264909, 264766, 18108376, 264692, 264693, 264694, 264695, 52644150, 264693, 11083381, 264909, 18108374, 35696423, 264634, 264557, 264638, 52644332, 83373044, 18108385, 56525486, 87168518, 22276902
2274	91610217 (4547, 4548)	Novel Protein sim. GBank gji1460112(emt)CA079611 - (X35942) HP1-SP74 protein [Mus musculus]	Contains protein domain (PF00538) linker histone H1 and H5 family	histone	52645156, 22276987, 22276989, 52645080, 29331824, 29331825, 29331826, 29331827, 29331828, 264905, 264908, 52644045, 264511, 265008, 265009, 80170831, 264591, 21906754, 33105954, 265011, 265018, 11083351, 264446, 264288, 264909, 264766, 18108376, 264692, 264693, 264694, 264695, 52644150, 264693, 11083381, 264909, 18108374, 35696423, 264634, 264557, 264638, 52644332, 83373044, 18108385, 56525486, 87168518, 22276902
2275	88062001 (4549, 4550)	Novel Protein sim. GBank gji13165408 (AC004755) - fos37502.2 [Homo sapiens]	Contains protein domain (PF00122) E1-E2 ATPase	transport	52645156, 22276987, 22276989, 52645080, 29331824, 29331825, 29331826, 29331827, 29331828, 264905, 264908, 52644045, 264511, 265008, 265009, 80170831, 264591, 21906754, 33105954, 265011, 265018, 11083351, 264446, 264288, 264909, 264766, 18108376, 264692, 264693, 264694, 264695, 52644150, 264693, 11083381, 264909, 18108374, 35696423, 264634, 264557, 264638, 52644332, 83373044, 18108385, 56525486, 87168518, 22276902
2278	11287447 (4551, 4552)			UNCLASSIFIED	264555, 264556

2284	95414825 (4567, 4568)	Novel Protein sim. GBank g 2498797 p 06431 IPNAO_MOUSE - PROTEIN N- TERMINAL ASPARAGINE AMIDOHYDROLASE (PROTEIN NH2-TERMINAL ASPARAGINE DEAMIDASE) (NTN- AMIDASE) (IPNAO) (PROTEIN NH2-TERMINAL ASPARAGINE AMIDOHYDROLASE) (IPNAO)			60424178, 35944507, 18100394, 52646842, 22278994, 35962696, 22278990, 22278997, 22278995, 264239, 60430409, 23031022, 23331824, 35950062, 20318395, 5264006, 5618245, 33657402, 60433438, 33100984, 5618245, 86589542, 87168559, 265018, 265019, 56811150, 264632, 264359, 21908785, 21908766, 21908787, 21908788, 21908769, 59511957, 35959917, 265020, 265021, 60170615, 33657023, 33657182, 27486262, 27486264, 27486265, 18100376, 55910784, 35698423, 35698855, 60431890, 87168518, 60432113, 254442, 265554
2285	87781484 (4569, 4570)	Novel Protein sim. GBank g 3342234 (U93369) - nuclear antigen EBNA-1 [Cervicopharynx herpesvirus 19]	collagen		35996052, 264005, 264907, 264908, 264909, 264512, 265003, 264910, 264595, 264700, 18100351, 264682, 264763, 264685, 264766, 264686, 264766, 264653, 264623, 35959555, 2646531, 2646535, 35959556, 35959535, 60170831, 384591, 35959520, 264692, 264593, 264594, 264595, 55912038, 264596, 87168474, 35695917, 264692, 55911578, 264555, 264557, 264692
2286	87737825 (4571, 4572)	Novel Protein sim. GBank g 3873414 (U00043) - similar to D. melanogaster fibrillar protein [Caenorhabditis elegans]	kinase		264631, 264633, 35959556, 35959535, 60170831, 384591, 35959520, 264692, 264593, 264594, 264595, 55912038, 264596, 87168474, 35695917, 264692, 55911578, 264555, 264557, 264692
2287	52956096 (4573, 4574)	Novel Protein sim. GBank g 630905 p S Q2131 - collagen [Homo sapiens]			
2288	94133083 (4575, 4576)	Novel Protein sim. GBank g 27883282 p P30189 ALU2_HUMAN - III ALU SUBFAMILY SB WARNING ENTRY III	UNCLASSIFIED		22278995, 56994075, 22278997, 22278999, 264259, 60432289, 264508, 264512, 265008, 33657402, 265017, 265018, 265019, 18100351, 264448, 21908765, 21908766, 21908767, 21908768, 33657023, 264457, 22279000, 22279002
2289	80094133 (4577, 4578)	Novel Protein sim. GBank g 2887487 (AC004144) - R34001_1 [Homo sapiens]	kinase		
2290	80084137 (4579, 4580)	Novel Protein sim. GBank g 2887487 (AC004144) - R34001_1 [Homo sapiens]	UNCLASSIFIED		264683
2291	84295281 (4581, 4582)	Novel Protein sim. GBank g 2853128 (AC005175) - R31449_3 [Homo sapiens]	shroud		18100394, 384907, 265008, 265009, 33100994, 52646317, 265010, 18100351, 264681, 264686, 264692, 18100370, 18100374, 18100385

2208	95312207 (4595, 4596)	Novel Protein sim. GBank g1307505 (tembic/CAR02849; ZN1050) predicted using GeneFinder, similar to collagen. cDNA EST EMBL D855564 comes from this gene, cDNA EST EMBL D69046 comes from this gene, cDNA EST y358b12.3 comes from this gene, cDNA EST y359b12.5 comes from this gene ...		collagen	60924170, 58181686, 22278959, 35966266, 22278996, 22278998, 22278999, 2644490, 264250, 29331822, 20331824, 66714117, 60424269, 35696052, 20331828, 66712502, 51682435, 33109510, 265006, 60433438, 21908754, 33109854, 55811386, 265010, 265018, 55811150, 264762, 18109351, 264682, 264683, 264288, 264684, 364686, 264688, 56181552, 284889, 21908788, 21908767, 29146626, 55811937, 21847894, 35695917, 265020, 18108362, 3567023, 18108394, 35691109, 55813702, 18108370, 60424269, 35696052, 20331828, 66714117, 35695945, 264630, 264634, 60431350, 18108390, 56182323, 264655, 83373044, 18108385, 60432113, 22276900, 264482, 264587, 264586
2209	80193720 (4597, 4598)	Novel Protein sim. GBank gi243386 (AC002294) - Unknown protein [Arabidopsis thaliana]		UNCLASSIFIED	2643369, 22278986, 22278989, 264259, 29331824, 66714117, 35696052, 264509, 264605, 254909, 264907, 264908, 265008, 264910, 265009, 264758, 265010, 87168559, 264600, 265018, 264760, 294762, 18108351, 264764, 264766, 264768, 294769, 21908768, 21908767, 35695917, 265021, 264636, 264638, 264639, 264631, 18108374, 35964223, 35695945, 264631, 264636, 264638, 18108385, 22276902, 264583
2301	91232575 (4601, 4602)	Novel Protein sim. GBank gi2143337 (p1)[4505 - calcium-dependent acetyl-binding protein - rat]		stud	264608, 264758, 265017, 21908765, 83373044, 264583
2302	86084141 (4603, 4604)	Novel Protein sim. GBank gi2837497 (AC004144) - R34001, 1 [Homo sapiens]		UNCLASSIFIED	52844045, 265019, 264286, 33657023, 18108370, 18108385
2303	94114139 (4605, 4606)	Novel Protein sim. GBank gi4884194 (emb)CA634220.1 - (AL049946) hypothetical protein [Homo sapiens]	- struct	Contains protein domain (PF00047) - Immunoglobulin domain	264259, 60430248, 264907, 265006, 264910, 6043229, 33657402, 265011, 265018, 264762, 264448, 264769, 264537, 264638, 83373044, 264486
2304	94304334 (4607, 4608)	Novel Protein sim. GBank gi2484162 (p1)Q10007Y1 - CAEL - HYPOTHETICAL (HYPOTHETICAL FISH: 1 IN CHROMOSOME II PRADUSO)		UNCLASSIFIED	264459, 29331824, 21908767, 3567182, 35673749
2305	90335911 (4609, 4610)	Novel Protein sim. GBank gi492566 (gi492566) - (AF132150) unknown (Drosophila melanogaster).			65214572, 22278998, 264608, 265008, 21908769, 264681, 264486

2300	65334940 (4611, 4612)	Novel Protein sim. GBank g14529265g1A033043.1 (AF15180) CGI-48 protein (Homo sapiens)	Contains protein domain (PF00400) WD domain, G-beta repeat	kinasereceptor	264488, 22278955, 22278956, 35696286, 22278957, 22278958, 22278959, 264298, 23331822, 23331824, 23331825, 23331826, 23331827, 23331828, 35696052, 264508, 264805, 264807, 23331830, 264808, 264809, 264510, 265006, 264511, 265007, 264512, 265008, 265009, 33657402, 21900754, 85658542, 265010, 265011, 264601, 265017, 265018, 264604, 265019, 19108351, 264448, 264288, 264766, 264769, 21900768, 29148629, 21900768, 21900787, 21900768, 29148629, 29148784, 35695917, 265020, 265021, 265022, 33637023, 264692, 16108370, 18108371, 18108376, 35698423, 35698555, 264838, 264839, 264840, 264841, 264842, 264838, 264839, 18108382, 18108385, 18108387, 264853, 264855, 264486
2307	79415283 (4613, 4614)	Novel Protein sim. GBank g14758732j1nFP_004522.1 (pmOCS - molybdenum cofactor synthesis 2)		UNCLASSIFIED	264828
2308	67608409 (4615, 4616)	Novel Protein sim. GBank g14758732j1nFP_004522.1 (pmOCS - molybdenum cofactor synthesis 2)		synthase	35696286, 264298, 26331832, 23331824, 264112, 264512, 264757, 21906754, 264288, 264690, 27486534, 264631, 264634, 264604, 18108397, 22278956, 22278957, 22278958, 22278959, 60432048, 26331822, 23331826, 60432289, 66712502, 60432229, 60433356, 60433438, 65274444, 265010, 264600, 264681, 264448, 294683, 264288, 21906766, 21906768, 265020, 264691, 264662, 264693, 65274920, 65274791
2309	95357218 (4617, 4618)	Novel Protein sim. GBank g13076055emb[CAB17070] - (Z99942) cDNA EST EMBL D73444 comes from this gene; cDNA EST EMBL D70905 comes from this gene; cDNA EST EMBL D72208 comes from this gene; cDNA EST EMB D72630 comes from this gene; cDNA EST EMB D72944 comes from this gene; cDN...		glycoprotein	264828, 22278955, 22278956, 35696286, 22278957, 22278958, 22278959, 264298, 23331822, 23331824, 23331825, 23331826, 264510, 264511, 265006, 264907, 35182433, 264510, 264511, 265006, 264512, 265007, 265008, 265009, 60433438, 33109954, 265010, 265011, 264603, 265017, 18108351, 264762, 264683, 264288, 264369, 264888, 33657023, 20281148, 20281089, 264628, 2633972, 55811576, 35696423, 20281071, 264632, 264636, 18108385, 18108387, 67168518, 22279000, 264553, 284486
2310	79601668 (4619, 4620)	Novel Protein sim. GBank g12137337j1n[C146281 - gene mCBP protein - mouse	Contains protein domain (PF00013) KH domain	UNCLASSIFIED	18108397, 55102575, 22278958, 56954075, 264298, 23331824, 23331837, 264596, 264907, 35182433, 264510, 264511, 265006, 264512, 265007, 265008, 265009, 60433438, 33109954, 265010, 265011, 264603, 265017, 18108351, 264762, 264683, 264288, 264369, 264888, 33657023, 20281148, 20281089, 264628, 2633972, 55811576, 35696423, 20281071, 264632, 264636, 18108385, 18108387, 67168518, 22279000, 264553, 284486

2317	87020571 (4633, 4634)		UNCLASSIFIED	22278998, 60432049, 264910, 60432729, 264686, 264687, 264688, 264689, 264558, 18108355
2318	7959879 (4635, 4636)		UNCLASSIFIED	264438, 264569, 18108336, 52646365,
2319	95101781 (4637, 4638)	Novel Protein sim. GBank gi 5262613 emb CA645746.1 - (AL080155) hypodermal protein [Homo sapiens]	UNCLASSIFIED	22278994, 22278995, 22278996, 56994075, 35696286, 22278997, 22278998, 2644259, 52645090, 29331825, 29331826, 29331827, 29331828, 29331830, 56182435, 60170831, 60432229, 60431735, 33657402, 21809794, 52644296, 87188474, 285011, 87168559, 265017, 265018, 265019, 18108351, 264448, 18108354, 264288, 264389, 32644229, 52645091, 18108352, 18108353, 21809795, 21809797, 21809798, 265002, 265003, 21809799, 21809800, 265001, 265002, 265003, 52644150, 33657023, 52645123, 30857109, 27486924, 33657249, 35695763, 18108370, 18108375, 18108379, 35696423, 264558, 83373944, 18108385, 56526486, 87168518, 264584, 264585, 264586
2320	91622456 (4639, 4640)	Novel Protein sim. GBank gi 728837 sp 9194ALU7. HUMAN - III ALU SUBFAMILY SO WARNING ENTRY III	kinase	22278994, 60432049, 60432789, 29331827, 264511, 265008, 52646317, 285017, 285019, 21806765, 18108372, 18108387, 22279002
2321	94320377 (4641, 4642)	Novel Protein sim. GBank gi 373837 emb CA027001 - (C0102) Similarity to S pombe hypodermal protein C.UDC9C (GNQ10194). cDNA EST EMBL110543 comes from the same tissue as EMBL101662 comes from this gene. cDNA EST EMBL101321 comes from this gene, cDNA EST EMBL102288 con....	UNCLASSIFIED	264488, 264687, 18108394, 264689, 21806765, 18108397, 18108398, 21809787, 21809788, 65274791, 22278995, 26595855, 22278998, 265021, 265022, 264510, 265006, 264511, 264512, 265008, 60170815, 264555, 60432230, 33657024, 264558, 60432231, 33657025, 264558, 264693, 60433356, 264559, 60433438, 264693, 60433356, 264559, 60433438, 29331824, 18108385, 18108386, 18108384, 29331825, 18108385, 33108954, 29331827, 29331828, 18108385, 33108954, 29331827, 56592486, 29146469, 265011, 60432113, 265017, 265018, 264508, 264563, 264482, 264509, 18108351, 264448, 284907, 284882, 18108370, 264683, 264908, 264286, 264609, 18108354, 264486, 264567
2322	87603165 (4643, 4644)	Novel Protein sim. GBank gi 507857 emb CA051665.1 - (AL109630) BACR7A4 y [Drosophila melanogaster]	Contains protein domain (PF00106) - dehydrogenase short chain dehydrogenase	22278996, 264708, 264511, 264737, 18108351, 264708, 264638

23223	9404043 (4645, 4646)	Novel Protein sim: GBank gi04916290 P1005 TRI_CAEEL_HYPOTHETICAL 59 kD PROTEIN T1105 H1 PRECURSOR	Cartilagin protein domain (PF00226)	eph	22278994, 22278995, 22278997, 60432049, 604259, 20331822, 22278997, 264509, 50182435, 264511, 285008, 60433356, 60433430, 55912038, 33109954, 2100754, 85855542, 87168474, 265011, 87168559, 265017, 265019, 264760, 264681, 18106351, 264369, 264288, 18106355, 264687, 264688, 21906765, 21906787, 21906768, 55811957, 35695917, 265021, 33657023, 18108382, 27486262, 55811576, 264631, 264555, 83373044, 87168518, 60432113, 22277002, 264592, 264593, 265020
23224	56635607 (4647, 4648)	Novel Protein sim: GBank gi04916290 P1005 TRI_CAEEL_HYPOTHETICAL 59 kD PROTEIN T1105 H1 PRECURSOR	Cartilagin protein domain (PF00226)	eph	22278994, 22278995, 22278997, 60432049, 604259, 20331822, 22278997, 264509, 50182435, 264511, 285008, 60433356, 60433430, 55912038, 33109954, 2100754, 85855542, 87168474, 265011, 87168559, 265017, 265019, 264760, 264681, 18106351, 264369, 264288, 18106355, 264687, 264688, 21906765, 21906787, 21906768, 55811957, 35695917, 265021, 33657023, 18108382, 27486262, 55811576, 264631, 264555, 83373044, 87168518, 60432113, 22277002, 264592, 264593, 265020
23225	88165074 (4649, 4650)	Novel Protein sim: GBank gi04916290 P1005 TRI_CAEEL_HYPOTHETICAL 59 kD PROTEIN T1105 H1 PRECURSOR	Cartilagin protein domain (PF00226)	eph	22278994, 22278995, 22278997, 60432049, 604259, 20331822, 22278997, 264509, 50182435, 264511, 285008, 60433356, 60433430, 55912038, 33109954, 2100754, 85855542, 87168474, 265011, 87168559, 265017, 265019, 264760, 264681, 18106351, 264369, 264288, 18106355, 264687, 264688, 21906765, 21906787, 21906768, 55811957, 35695917, 265021, 33657023, 18108382, 27486262, 55811576, 264631, 264555, 83373044, 87168518, 60432113, 22277002, 264592, 264593, 265020
23226	84390062 (4651, 4652)	Novel Protein sim: GBank gi04916290 P1005 TRI_CAEEL_HYPOTHETICAL 59 kD PROTEIN T1105 H1 PRECURSOR	Cartilagin protein domain (PF00226)	eph	22278994, 22278995, 22278997, 60432049, 604259, 20331822, 22278997, 264509, 50182435, 264511, 285008, 60433356, 60433430, 55912038, 33109954, 2100754, 85855542, 87168474, 265011, 87168559, 265017, 265019, 264760, 264681, 18106351, 264369, 264288, 18106355, 264687, 264688, 21906765, 21906787, 21906768, 55811957, 35695917, 265021, 33657023, 18108382, 27486262, 55811576, 264631, 264555, 83373044, 87168518, 60432113, 22277002, 264592, 264593, 265020
23227	88081648 (4653, 4654)	Novel Protein sim: GBank gi04916290 P1005 TRI_CAEEL_HYPOTHETICAL 59 kD PROTEIN T1105 H1 PRECURSOR	Cartilagin protein domain (PF00226)	eph	22278994, 22278995, 22278997, 60432049, 604259, 20331822, 22278997, 264509, 50182435, 264511, 285008, 60433356, 60433430, 55912038, 33109954, 2100754, 85855542, 87168474, 265011, 87168559, 265017, 265019, 264760, 264681, 18106351, 264369, 264288, 18106355, 264687, 264688, 21906765, 21906787, 21906768, 55811957, 35695917, 265021, 33657023, 18108382, 27486262, 55811576, 264631, 264555, 83373044, 87168518, 60432113, 22277002, 264592, 264593, 265020
23228	83388428 (4655, 4656)	Novel Protein sim: GBank gi04916290 P1005 TRI_CAEEL_HYPOTHETICAL 59 kD PROTEIN T1105 H1 PRECURSOR	Cartilagin protein domain (PF00226)	eph	22278994, 22278995, 22278997, 60432049, 604259, 20331822, 22278997, 264509, 50182435, 264511, 285008, 60433356, 60433430, 55912038, 33109954, 2100754, 85855542, 87168474, 265011, 87168559, 265017, 265019, 264760, 264681, 18106351, 264369, 264288, 18106355, 264687, 264688, 21906765, 21906787, 21906768, 55811957, 35695917, 265021, 33657023, 18108382, 27486262, 55811576, 264631, 264555, 83373044, 87168518, 60432113, 22277002, 264592, 264593, 265020
23229	87604478 (4657, 4658)	Novel Protein sim: GBank gi04916290 P1005 TRI_CAEEL_HYPOTHETICAL 59 kD PROTEIN T1105 H1 PRECURSOR	Cartilagin protein domain (PF00226)	eph	22278994, 22278995, 22278997, 60432049, 604259, 20331822, 22278997, 264509, 50182435, 264511, 285008, 60433356, 60433430, 55912038, 33109954, 2100754, 85855542, 87168474, 265011, 87168559, 265017, 265019, 264760, 264681, 18106351, 264369, 264288, 18106355, 264687, 264688, 21906765, 21906787, 21906768, 55811957, 35695917, 265021, 33657023, 18108382, 27486262, 55811576, 264631, 264555, 83373044, 87168518, 60432113,

2335	60046102 (4668, 4670)	Novel Protein sim. GBank gll283330 (AF062378) - calmodulin-binding protein SH41 [Mus musculus]	Contains protein domain (PF00612) - Ruvcl	18108351, 21000769, 264555
2336	95196121 (4671, 4672)	Novel Protein sim. GBank gll929056 (mjcA72805) - (Y12090) putative 3,4-dihydroxy-2-butanone kinase [Lycopodium obscurum]	kinase	264907, 35693917, 18108379
2337	95345810 (4673, 4674)	Novel Protein sim. GBank gll489063 (mjcCA939181.1) - (Z65998) dJ108K11.3 (similar to yeast suppressor protein SRP40) [Homo sapiens]	UNCLASSIFIED	35696286, 22278999, 56102181, 29331825, 6042489, 56182435, 33974702, 55810038, 55811386, 265017, 265018, 265019, 21900768, 35699917, 264991, 33857023, 33857109, 263972, 35696423, 35696655, 60432113
2338	87634045 (4675, 4676)	Novel Protein sim. GBank gll224688 (mjcCA939181.1) - (AB002372) KIA0374 [Homo sapiens]	UNCLASSIFIED	18108394, 29331822, 60714117, 60432280, 264596, 265010, 265019, 18108364, 6042438, 264369, 55811957, 265021, 33657023, 263976, 55811576, 264632, 56182233, 264639
2339	85065319 (4677, 4678)	Novel Protein sim. GBank gll387355 (mjcCA939181.1) - (AL033834) serine-rich protein [Schistosoma mansoni]	UNCLASSIFIED	35696286, 264592, 264367, 264991, 264653
2340	90937716 (4679, 4680)			65274572, 22278994, 35696786, 22278997, 22278999, 264259, 29331822, 60432289, 29331826, 29331830, 265006, 33857402, 33109954, 265017, 265018, 264766, 264685, 21900769, 35699917, 264991, 264692, 35696423, 87168519, 22279000
2341	87715381 (4681, 4682)	Novel Protein sim. GBank gll387355 (mjcCA939181.1) - (Z31042) similar to Yeast cytoplasmic protein YEF8 like; cDNA EST Y208B4.3 comes from this gene; cDNA EST Y420B4.5 comes from this gene; cDNA EST Y4303n1.3 comes from this gene; cDNA EST Y4303n1.5 comes from this gene; cDNA ...	Contains protein domain (PF00400) - kinase WD domain, C-leuc repeat	264486, 65274572, 22278997, 22278999, 264259, 29331822, 60432289, 29331826, 29331830, 265006, 33857402, 66714117, 29331825, 60432289, 35696052, 264005, 264599, 29331830, 265008, 2670831, 60432229, 60433356, 87168474, 265017, 265018, 265019, 264448, 264398, 264383, 21900768, 21903788, 55811957, 35695917, 265020, 265022, 52644150, 33857023, 65274620, 33657109, 18108370, 18108376, 55810764, 35696423, 55811576, 264596, 264658, 18108385, 60432113, 264583, 264594, 264595, 264596, 264597
2343	87755448 (4685, 4686)	Novel Protein sim. GBank gll462971 (mjcCA939181.1) - (AF151894) CG1-136 protein [Homo sapiens]	UNCLASSIFIED	264907, 264512, 265011, 264683
2344	79553191 (4687, 4688)	Novel Protein sim. GBank gll260507 (mjcCA939181.1) - (AF151894) CG1-136 protein [Homo sapiens]	UNCLASSIFIED	264758
2345	94119795 (4689, 4690)	Novel Protein sim. GBank gll260507 (mjcCA939181.1) - (AF151894) CG1-136 protein [Homo sapiens]	Contains protein domain (PF00092) - collagen von Willebrand factor type A domain	264486, 264259, 66714502, 264756, 9337304, 264566

2346	84131820 (4691, 4692)	Novel Protein sim. G.Bank g11725411 (U53153) - one short region of weak similarity to <i>S. cerevisiae</i> protease A inhibitor 3 (SP-P01084) and another short region of weak similarity to <i>S. cerevisiae</i> glucose repression mediator protein (SP-P14922) (<i>Caenorhabditis elegans</i>)	Contains protein domain (PF00515) - proteaseinhib PR Domain	33596286, 22278998, 264259, 35960052, 26331828, 33657402, 60433366, 33109954, 87168559, 264603, 265019, 18108351, 264681, 264685, 21906786, 265021, 33057109, 55811576, 35695855, 264637, 52644332, 264597, 83373044, 22279000, 22279002
2347	85330387 (4693, 4694)			22278997, 264511, 264683, 264684, 264768, 264687, 264688, 264691, 264692, 55811576, 18108351, 35695855, 264693, 264697, 264698, 264699, 264700, 264907, 264908, 264909, 264516, 264511, 265008, 264907, 264512, 265008, 265009, 264910, 264991, 264592, 264593, 264694, 264737, 264595, 264596, 264758, 265011, 264691, 264762, 18108351, 264764, 264288, 264766, 264768, 264889, 35695917, 264693, 264628, 18108370, 264629, 18108374, 35696423, 264631, 264635, 264636, 264637, 264638, 264639, 83373044, 18108385, 264597, 264488
2349	87776502 (4697, 4698)	Novel Protein sim. G.Bank g14884106(jem)CAB324 (I - (ALD0562) hypothetical protein (Homo sapiens))		35696052, 29146499, 264909, 264369
2350	88260394 (4698, 4700)			22278998, 22278999, 264259, 26331822, 26331824, 29331825, 29331827, 29331828, 33109954, 21906784, 265010, 87168559, 264681, 264685, 21906786, 265021, 18108351, 35695855, 264637, 52644332, 264597, 83373044, 22279000, 22279002, 264482
2351	86568042 (4701, 4702)	Novel Protein sim. G.Bank g11728832(lgIP3876)HUMAN - III ALU SUBFAMILY SB WARNINGS ENTRY III	Kinase	551182575, 264909, 265008, 264558
2352	87337196 (4703, 4704)	Novel Protein sim. G.Bank g1171637(lgIP3876)YH45 - HYPOTHETICAL 75.9 KD PROTEIN IN SPO13-ARGA INTERGENIC REGION	Contains protein domain (PF00076) - UNCLASSIFIED RNA recognition motif (a.k.a. RRM, RBD, or RNP domain)	264259, 264448
2353	91638784 (4705, 4706)	Novel Protein sim. G.Bank g11489595(lgIP4889)RR2 - DROME - HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN 27C (HNRNP 48) (HPR48.1)	Contains protein domain (PF00076) - RNA recognition motif (a.k.a. RRM, RBD, or RNP domain)	29331826, 55812038, 265019, 264692, 264636
2354	87337196 (4707, 4708)	Novel Protein sim. G.Bank g11489595(lgIP4889)RR2 - DROME - HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN 27C (HNRNP 48) (HPR48.1)	Contains protein domain (PF00076) - RNA recognition motif (a.k.a. RRM, RBD, or RNP domain)	29331824, 264908, 265008, 265008

2362	91721193 (473, 473a)	Novel Protein sm. GBank gll117108330bp/197819WSB ACACA - MYOSIN HEAVY CHAIN IB [MYOSIN HEAVY CHAIN I.]		UNCLASSIFIED	22278959, 264259, 29331822, 29331824, 60432289, 264509, 264512, 60432229, 60433356, 264448, 264682, 264683, 264369, 21906765, 21906766, 21906769, 60432113, 22279000, 22279002, 264635
2363	95006535 (4725, 4725b)	Novel Protein sm. GBank gll854065[em]CAA83371- (X53413) U88 [Human herpesvirus 6]		UNCLASSIFIED	264458, 18108394, 264687, 18106398, 22278999, 59894075, 35069626, 22789897, 22278999, 264259, 60714117, 29331825, 35050052, 264509, 264605, 264506, 264507, 264508, 264509, 264510, 264511, 264512, 264513, 60714302, 264609, 264511, 264505, 264506, 264507, 264508, 264509, 21906764, 87168474, 265010, 87168559, 264603, 265017, 265018, 265019, 264760, 264762, 18108351, 264448, 264754, 264683, 264684, 264288, 18108355, 264768, 18108356, 264689, 18108359, 21906765, 21906766, 21906767, 35695917, 265020, 265021, 265022, 60170615, 52644150, 264691, 33657023, 264692, 18108394, 33657109, 18108368, 18108370, 18108374, 35695423, 35695955, 264635, 264556, 264557, 264639, 60170394, 83373044, 18108383, 18108394, 18106385, 18108388, 22278998, 22278999, 22278999, 264507, 264609, 264610, 33657402, 264758, 264600, 264766, 264687, 264689, 21906765, 264766, 264687, 264689, 21906769, 265021, 33657023, 33657109, 83373044, 264566
2364	94140746 (4729, 4730)	Novel Protein sm. GBank gll140045 (U49382) - transporter protein [Homo sapiens]		transport	52644507, 264259, 29331824, 29331825, 29331826, 29331827, 29331828, 264907, 29331830, 264909, 264511, 265008, 33657402, 264595, 52946317, 265017, 265018, 265019, 264605, 264685, 264766, 264689, 21906766, 21906769, 35695917, 265020, 265021, 265022, 52644150, 35695955, 52644332, 18106385, 18108387, 264554, 264556, 264557, 21906768, 55810764, 65274791, 264629
2365	94140746 (4729, 4730)	Novel Protein sm. GBank gll140045 (U49382) - transporter protein [Homo sapiens]		transport	52644507, 264259, 29331824, 29331825, 29331826, 29331827, 29331828, 264907, 29331830, 264909, 264511, 265008, 33657402, 264595, 52946317, 265017, 265018, 265019, 264605, 264685, 264766, 264689, 21906766, 21906769, 35695917, 265020, 265021, 265022, 52644150, 35695955, 52644332, 18106385, 18108387, 264554, 264556, 264557, 21906768, 55810764, 65274791, 264629
2366	94312388 (4731, 4732)	Novel Protein sm. GBank gll166451 (U40110) - C56G1.4 gene product [Camponotus degani]		UNCLASSIFIED	264554, 264556, 264557, 21906768, 55810764, 65274791, 264629
2367	941406910 (4733, 4734)	Novel Protein sm. GBank gll166451 (U40110) - C56G1.4 gene product [Camponotus degani]		UNCLASSIFIED	264554, 264556, 264557, 21906768, 55810764, 65274791, 264629
2368	94321210 (4735, 4736)	Novel Protein sm. GBank gll166451 (U40110) - C56G1.4 gene product [Camponotus degani]		UNCLASSIFIED	264554, 264556, 264557, 21906768, 55810764, 65274791, 264629

2369	9431334 (4737, 4738) (AB028343) a-helix coiled-coil rod homologue [Homo sapiens]	Novel Protein sim. GBank gi 5360907 db BAA482158.1 - (AB028343) a-helix coiled-coil rod homologue [Homo sapiens]	52644507, 52648842, 35696286, 264092, 264094, 52645090, 35696054, 264107, 29331830, 52644045, 265006, 265007, 265009, 52644296, 52644220, 264686, 21906765, 21906766, 35695917, 265020, 52644150, 263967, 33657106, 27468265, 35695783, 18108370, 203974, 18108374, 18108376, 52644332, 263981, 18108385, 264505, 264906, 264936	stud
2370	75804120 (4739, 4740)		UNCLASSIFIED	
2371	57264006 (4741, 4742)		UNCLASSIFIED	
2372	87642413 (4743, 4744)		UNCLASSIFIED	
2373	87418611 (4745, 4746)	Novel Protein sim. GBank gi 4589532 db BAA76813.1 - (AB023166) KIAA0959 protein [Homo sapiens]	26331136, 265010, 265018, 35695917, 264634, 60432113	UNCLASSIFIED
2374	94123865 (4747, 4748)	Novel Protein sim. GBank gi 5103131 db BAA80445.1 - (AP000081) 246aa long hypothetical ribonuclease PH [Aeropyrum pernix]	265006, 265007, 265008, 265009, 265011, 264766, 35695917, 35695865, 263981, 264557, 264365	UNCLASSIFIED
2375	87313352 (4749, 4750)	Novel Protein sim. GBank gi 13511196 P47758 SRP5, MOUSE - SIGNAL RECOGNITION PARTICLE RECEPTOR BETA SUBUNIT (SR-BETA)	60432046, 20331824, 264097, 52644045, 264512, 60433338, 21906754, 5264296, 87185559, 264448, 21906765, 21906768, 21906769, 33657023, 18108368, 55811576, 52644332	UNCLASSIFIED
2376	87613744 (4751, 4752)	Novel Protein sim. GBank gi 2645435 (AF007807) - CHQ3 [Chroococcia melanogaster]	264259, 26331830, 265009, 264910, 265006, 60433438, 21906754, 265017, 265018, 265019, 264682, 264288, 264685, 21906767, 263972, 35695935, 87188518, 60432113	ATPase-associated PRD-finger
2377	95316689 (4753, 4754)	Novel Protein sim. GBank gi 5270599 gi A041238.1 - (AF083249) Rb binding protein homolog [Homo sapiens]	18108394, 65274572, 22278697, 22278999, 265018, 265019, 265020, 265021, 6674117, 265022, 265023, 265024, 33656970, 20146496, 20146498, 26445026, 265006, 265007, 265008, 265009, 60170831, 265010, 265011, 265018, 55811150, 18108351, 264764, 264288, 21906787, 21906788, 20148627, 20148629, 265021, 33657023, 33657106, 18108370, 18108374, 18108378, 35696423, 264556, 83372044, 18108385, 18108386, 56558466, 22279000, 22279002, 264563	UNCLASSIFIED
2378	94137032 (4755, 4756)	Novel Protein sim. GBank gi 1072188 (U49642) - No definition the found [Caenorhabditis elegans]	65274572, 56182573, 35695866, 264259, 29331832, 26331824, 66741117, 29331825, 29331826, 60432286, 29331827, 56182435, 264510, 265009, 60433356, 97189474, 265011, 265018, 264288, 21906765, 18108363, 26395597, 56182432, 83373044, 18108385, 22279002, 22279003	UNCLASSIFIED
2379	85444322 (4757, 4758)	Novel Protein sim. GBank gi 1337357 (AC004481) - hypothetical protein [Arabidopsis thaliana]	265017, 264288, 21906766	UNCLASSIFIED

2387	14997990 (4773, 4774)			UNCLASSIFIED	264634
2388	11424604 (4775, 4776)			UNCLASSIFIED	264595
2389	95310550 (4777, 4778)	Novel Protein sim. GBank g14758029p[HP_004372.1]PCREB -cAMP responsive element binding protein-like 1	Contains protein domain (PF001070) - bZIP transcription factor	UNCLASSIFIED dna_rna_bind	264488, 22278998, 22278999, 264508, 264602, 264506, 222907, 264508, 264609, 265008, 264511, 264512, 264810, 264591, 21903754, 264601, 356904, 264768, 18108374, 264602, 264603, 264768, 264769, 21906765, 21908768, 264602, 264603, 35696423, 264635, 264636, 264555, 83373944, 22278000, 264486
2390	94320912 (4779, 4780)	Novel Protein sim. GBank g11644233[db][BAA12223] - [D04103] mitochondrial DNA polymerase gamma [Homo sapiens]	Contains protein domain (PF00478) - DNA polymerase family A	polymerase	52844507, 55182575, 22278993, 35698286, 22278998, 22278997, 22278998, 29331822, 29331825, 29331826, 35695052, 264605, 52844045, 265009, 264758, 264759, 33109954, 52844286, 85558542, 265011, 265017, 265018, 264605, 52844229, 21906765, 21906767, 21906768, 21906769, 35695917, 52844150, 33557023, 33657109, 33657349, 35695783, 18108370, 18108374, 18108376, 35696423, 35695855, 264555, 52844332, 55182323, 60170394, 83373044, 55292486
2391	60035194 (4781, 4782)			UNCLASSIFIED	264595
2392	94245016 (4783, 4784)	Novel Protein sim. GBank g14240169[db][BAA14863.1] - [A02020647] KIAA0840 protein [Homo sapiens]	Contains protein domain (PF003550) - Leucine Rich Repeat	Leucine-rich repeat	264598, 35695052, 264508, 264509, 264598, 264606, 264608, 264609, 264510, 264511, 264512, 264810, 265009, 264591, 264758, 264600, 264604, 264762, 264448, 264764, 264369, 264766, 264768, 264769, 264689, 35695917, 264629, 18108374, 263978, 35696423, 35695855, 264631, 264634, 264635, 264636, 264637, 264638, 60170394, 264639, 264585, 264488

2393	95302833 (4785, 4788) Novel Protein sim. GBank g M566837 refnp_000893.1 PRP - ribosomal protein, large, P0	Contains protein domain (PF00466) Ribosomal protein L10	18100392, 60424179, 264489, 18100394, 18100397, 22278995, 56594075, 35866286, 22279986, 22278997, 22278999, 264093, 60432049, 264259, 26331822, 29147820, 20281099, 29331824, 29331825, 26714117, 60432289, 29331826, 29331827, 29331828, 35696052, 29146499, 264500, 264509, 264905, 264907, 264908, 66712502, 52644045, 264628, 264909, 56182435, 284112, 264113, 264510, 265006, 264511, 265007, 265008, 265009, 294910, 264991, 264993, 60433386, 21906765, 5811386, 264994, 265010, 265011, 21906764, 5811386, 265010, 265011, 265017, 265018, 265019, 264681, 264762, 18100351, 264763, 264682, 264764, 264683, 264369, 264285, 18100354, 264768, 264686, 264687, 264688, 264689, 18100359, 21906765, 21906766, 21906767, 21906768, 21906769, 5811957, 29148629, 29148784, 35695917, 265021, 265022, 33657023, 264692, 264693, 18100384, 33657109, 18100388, 27486281, 27486282, 33657349, 35695763, 18100370, 263972, 264629, 18100374, 263977, 18100376, 263978, 5810764, 35696423, 35695855, 264634, 60431850, 264553, 264637, 264557, 263981, 264559, 18100381, 5070394, 35696286, 22278996, 264692, 5811386, 265010, 265011, 265017, 265018, 264448, 264683, 264288, 21906765, 21906768, 21906769, 5811957, 35695917, 33657023, 65274620, 33657182, 33657349, 35695763, 18100374, 18100376, 5810764, 5811576, 35696423, 60170394, 18100385, 264554, 264556, 264567
2394	94323266 (4787, 4788) Novel Protein sim. GBank g I159888 (AC004908) - zinc finger protein from gene of uncertain exon structure, similar to C99678 (P1D302533) (Homo sapiens)	Contains protein domain (PF00086) - dna_rna_bind Zinc finger, C2H2 type	35696286, 22278996, 264692, 5811386, 265010, 265011, 265017, 265018, 264448, 264683, 264288, 21906765, 21906768, 21906769, 5811957, 35695917, 33657023, 65274620, 33657182, 33657349, 35695763, 18100374, 18100376, 5810764, 5811576, 35696423, 60170394, 18100385, 264554, 264556, 264567
2395	95287212 (4789, 4790) Novel Protein sim. GBank g J1712755 A047636.1 AF16079 - (AF16079) calcium transporter Cat1 [Rattus norvegicus]	dna_rna_bind	264259, 29331824, 264910, 264288, 265021, 83373044, 18100387, 264663, 264666

2396	95096700 (4791, 4792)	Novel Protein sim. GBank g[106322]p[1514037] - hypothetical protein (LH 3' region) - human	Contains protein domain (PF00560) - nuclease Leucine Rich Repeat	52646365, 18106397, 5612575, 35696286, 22278997, 22278998, 60432049, 264259, 29331822, 29331824, 29331825, 29331826, 29331827, 29331828, 264106, 264907, 29331830, 66712502, 264110, 60170831, 264391, 33657402, 60433433, 55812038, 33109654, 21906754, 33657084, 87168474, 265017, 265018, 265019, 264760, 264448, 264288, 264766, 52644225, 21906766, 21906767, 265020, 265021, 60170815, 264692, 33657023, 85274620, 52645129, 33657102, 27486267, 27486268, 27486269, 264683, 18106397, 35696286, 264907, 264683, 18106397, 35696286, 264907, 85373044, 18106398, 87168518, 22279002, 264482
2397	87280354 (4793, 4794)			52644507, 52645155, 5612575, 264259, 29147620, 264905, 264907, 264908, 264609, 264910, 264758, 52644296, 264603, 264604, 264782, 264681, 264794, 18106357, 264769, 21906768, 264693, 264628, 264635, 264638, 264639, 264584
2398	86047659 (4795, 4796)	Novel Protein sim. GBank g[12358609] (AC05179) - H83 GST [Homo sapiens]	UNCLASSIFIED	
2399	87738965 (4797, 4798)	Novel Protein sim. GBank g[1796117] (L41394) - nuclear protein [Ensis minor]	UNCLASSIFIED	56994075, 22278997, 22278998, 22278999, 264259, 29331822, 29331824, 60714117, 60432289, 264906, 29331830, 56182435, 264112, 264910, 33109894, 21906794, 264683, 18106397, 35696286, 265019, 264784, 264685, 265015, 21906767, 21906767, 21906769, 35695917, 265020, 265022, 60170915, 33657023, 18106370, 18106374, 264556, 60170394, 264558, 18106374, 264556, 60170394, 264558, 87168518, 22279000, 22279002, 264594, 264566, 264457
2400	81214116 (4799, 4800)	Novel Protein sim. GBank g[1352822]p[AB69285.1] - (AF008943) glucose-6-phosphatase [Haplochromis nudioid]	phosphatase	21906766, 52646842, 56994075, 33657182, 27486282, 52644296, 265017

2401	91214118 (4801, 4802)	Novel Protein sim. GBank gl 2352822 p AA-B69285.1 - (AF008949) glucose-6-phosphatase [Haplochromis rubellus]	phosphatase	52644507, 52643156, 52644220, 264668, 21906764, 21906765, 52646365, 52646842, 21906766, 21906767, 21906768, 22278995, 35659917, 56994075, 35656286, 22278996, 22278997, 265020, 22278998, 22278999, 264439, 33657023, 52645080, 264693, 29331824, 33657109, 52645129, 29331826, 33657162, 29331827, 35696052, 27486261, 33657162, 35696052, 33657163, 27486265, 52645129, 264693, 29331827, 35696052, 35655885, 265006, 265007, 265008, 265009, 264637, 52644332, 56812038, 52646317, 18108335, 52644296, 87158474, 265010, 87188559, 50432113, 265017, 265018, 265019, 264553, 264288, 264807, 264908, 264909, 264566
2402	91221408 (4803, 4804)	Novel Protein sim. GBank gl 4689259 p AAD27832.1 pF12185. - (AF121859) sorting nexin 9 [Homo sapiens]		
2403	94135432 (4805, 4806)	Novel Protein sim. GBank gl 4925575 p AAD54048.1 pF15181. - (AF151811) CGI-53 protein [Homo sapiens]	Contains protein domain (PF00062) - C-type lysozyme/alpha-lactalbumin family	22278999, 35696052, 265018, 264886, 264693, 83373044, 264567
2404	55312602 (4807, 4808)	Novel Protein sim. GBank gl 2315735 (AF156689) - similar to short chain-type dehydrogenase [Caenorhabditis elegans]	Contains protein domain (PF00106) - short chain dehydrogenase	35656286, 29331828, 35696052, 265008, 265018, 21906769, 264564
2405	94311851 (4809, 4810)	Novel Protein sim. GBank gl 464176 p p A403581 - (D14853) polyprotein [Hepatitis C virus]	UNCLASSIFIED	35656286, 29331827, 265007, 21906764, 265017, 265018, 265019, 264763, 264369, 21906765, 35696017, 265020, 265021, 52644150, 264693, 35695955, 264632,
2406	88094501 (4811, 4812)	Novel Protein sim. GBank gl 2773383 (AF041382) - microtubule binding protein D-CLIP-190 [Drosophila melanogaster]	Contains protein domain (PF01302) - struct CAP-Gly domain	52646842, 22278994, 22278996, 35696286, 22278997, 22278998, 22278999, 264092, 264093, 60432049, 264259, 29331822, 29331824, 29331825, 29331827, 35696052, 29331828, 264102, 264106, 264906, 52644045, 265007, 265008, 265009, 264910, 264592, 60433356, 60433436, 33109954, 265010, 265011, 265018, 265019, 264369, 264485, 264686, 21906766, 21906769, 35696432, 264632, 36182223, 264638, 264765, 264667, 264700, 264953
2407	79465005 (4813, 4814)		UNCLASSIFIED	264410, 265010, 264448, 364557
2408	87391502 (4815, 4816)	Novel Protein sim. GBank gl 423442 p p J33513 - gene F1f protein - mouse	UNCLASSIFIED	

2409	8441770 (4817, 4818)	Novel Protein sim. GBank glt1176801bip4595fyn25 CAEL - HYPOTHETICAL 20.8 KD PROTEIN 108A3.6 IN CHROMOSOME III	UNCLASSIFIED	22278955, 22278956, 22278957, 264097, 264098, 264099, 264100, 264101, 52641822, 26331824, 29331827, 29314698, 52640405, 60433438, 33657084, 87168474, 264780, 21906767, 29148627, 29148628, 52644150, 33657023, 263967, 20281069, 19108374, 20281071, 59182323, 83373044, 19108385, 87168518 22278998, 264259
2410	8750460 (4619, 4620)	Novel Protein sim. GBank gll959522ipJAC48052.2) - (UC4849) Contains similarity to Plant domain; PF00546 (P- box). Score=28.17, E-value=4.3e-05, N=1 [Caenobabditia sp.]	Contains protein domain (PF00060) - Copper/zinc superoxide dismutase (SODC)	22278955, 22278956, 22278957, 264097, 264098, 264099, 264100, 264101, 52641822, 26331824, 29331827, 29314698, 52640405, 60433438, 33657084, 87168474, 264780, 21906767, 29148627, 29148628, 52644150, 33657023, 263967, 20281069, 19108374, 20281071, 59182323, 83373044, 19108385, 87168518 22278998, 264259
2411	87334632 (4821, 4822)	Novel Protein sim. GBank gll3114713 (PF061346) - Eup1 protein [Mus musculus]	trf	26331824, 29331827, 29331828, 264097, 264098, 264099, 264100, 264101, 52641822, 26331824, 29331827, 29314698, 52640405, 60433438, 33657084, 87168474, 264780, 21906767, 29148627, 29148628, 52644150, 33657023, 263967, 20281069, 19108374, 20281071, 59182323, 83373044, 19108385, 87168518, 22279002
2412	87778332 (4823, 4824)	Novel Protein sim. GBank gll410339ipJAC43038.1) - [AF106585] myelin gene expression factor 2 [Homo sapiens]	Contains protein domain (PF00076) - RNA recognition motif (a.k.a. RRM, RBD, or RNP domain)	26331824, 29331827, 29331828, 264097, 264098, 264099, 264100, 264101, 52641822, 26331824, 29331827, 29314698, 52640405, 60433438, 33657084, 87168474, 264780, 21906767, 29148627, 29148628, 52644150, 33657023, 263967, 20281069, 19108374, 20281071, 59182323, 83373044, 19108385, 87168518, 22279002
2413	84133820 (4825, 4826)	Novel Protein sim. GBank gll262705ipJAC434778.1) - (AL080214) hypothetical protein [Homo sapiens]	Contains protein domain (PF00038) - Intermediate filament proteins	26331824, 29331827, 29331828, 264097, 264098, 264099, 264100, 264101, 52641822, 26331824, 29331827, 29314698, 52640405, 60433438, 33657084, 87168474, 264780, 21906767, 29148627, 29148628, 52644150, 33657023, 263967, 20281069, 19108374, 20281071, 59182323, 83373044, 19108385, 87168518, 22279002
2414	84312590 (4827, 4828)	Novel Protein sim. GBank gll1082340ipJ552653 - DNA- binding protein K kappa B - human	ubiquitin	26331824, 29331827, 29331828, 264097, 264098, 264099, 264100, 264101, 52641822, 26331824, 29331827, 29314698, 52640405, 60433438, 33657084, 87168474, 264780, 21906767, 29148627, 29148628, 52644150, 33657023, 263967, 20281069, 19108374, 20281071, 59182323, 83373044, 19108385, 87168518, 22279002

2415	85098002 (4828, 4830)	Novel Protein sim. GBank g1423915 p I44538 - myosin I heavy chain - rat	Contains protein domain (PF00063) - Myosin head (motor domain)	struct	254259, 264908, 60433358, 33657402, 21900754, 2650118, 264687, 264689, 21900769, 55811957, 265021, 264680, 264681, 33657023, 264683, 35696422, 55182323, 56526486
2416	94118356 (4831, 4832)	Novel Protein sim. GBank g13025445 (AC004328) - R32184.1 [Homo sapiens]	Contains protein domain (PF00153) - Mitochondrial carrier proteins	struct	264084, 29331822, 29331824, 29331827, 264369
2417	87733334 (4833, 4834)	Novel Protein sim. GBank g11084944 p I554495 - hypothetical protein YP0021c - yeast (Saccharomyces cerevisiae)	Contains protein domain (PF00411) - Ribosomal protein S11	UNCLASSIFIED	58900755, 264091, 264259, 29331824, 29331825, 60432289, 29331828, 264605, 264807, 264511, 265009, 60432279, 21900754, 87186559, 265019, 264682, 21900768, 21906769, 265020, 265021, 33657023, 55274620, 18108370, 55811576, 264634, 80170394, 18108385, 22279000, 22279002, 264566
2419	82374249 (4837, 4838)	Novel Protein sim. GBank g1284000 p I518732 - autoantigen, 64K - human		struct	264569, 264762, 264448, 264681, 264631, 264634, 264555, 264560, 264638, 264558
2420	94944244 (4839, 4840)	Novel Protein sim. GBank g11070821 p I550755 - hypothetical protein VSP-3 - Chlamydomonas reinhardtii		UNCLASSIFIED	29331824, 29331825, 29331828, 60432229, 33109954, 86655542, 87168474, 265018, 264288, 265020, 264554
2421	87805345 (4841, 4842)	Novel Protein sim. GBank g12245957 p I5AA207721 - (AB02311) KIAA0313 [Homo sapiens]	Contains protein domain (PF00617) - RasGEF domain	UNCLASSIFIED	264309, 264768, 264539
2422	88084714 (4843, 4844)	Novel Protein sim. GBank g12245957 p I5AA207721 - (AB02311) KIAA0313 [Homo sapiens]	Contains protein domain (PF00617) - RasGEF domain	transport	18108392, 18108394, 18108398, 264606, 265008, 265010, 18108391, 18108394, 18108395
2423	88058390 (4845, 4846)	Novel Protein sim. GBank g14305153 p I5A_022892.1 p I5A0X - MAPK kinase kinase 3	Contains protein domain (PF00058) - Eukaryotic protein kinase domain	kinase	264259, 60432049, 29331822, 29331826, 60432289, 29331828, 265008, 265009, 60433356, 21906754, 265017, 265018, 265019, 21906766, 21906768, 21906769, 265020, 265021, 20281148, 263971, 60432113
2424	84854047 (4847, 4848)	Novel Protein sim. GBank g12888398 (AC004381) - Unknown gene product [Homo sapiens]		UNCLASSIFIED	55182575, 35596286, 22278997, 60432049, 264259, 29331826, 29331828, 264605, 66712502, 29331830, 60433356, 265011, 265019, 264766, 21906766, 55811957, 264682, 33657023, 33657109, 55811576, 56182323, 83373044, 18108385, 18108389, 60432113, 22279000
2425	87415881 (4849, 4850)	Novel Protein sim. GBank g1007752 p I5A198761 - (D8F558) Protein Kinase [Rattus norvegicus]	Contains protein domain (PF00069) - kinase	kinase	22278998, 22278998, 264259, 264102, 264512, 265008, 21906767, 18108370, 18108374, 263976
2426	87613945 (4851, 4852)	Novel Protein sim. GBank g1003636 p I5A35003.1 - (U94619) circulating cathodic antigen [Schistosoma mansoni]	Eukaryotic protein kinase domain	UNCLASSIFIED	

2427	87622692 (4853, 4854)	Novel Protein sim. GBank gi4680693/gpAAD2737 (AF13296 - (AF132962) CGI-26 protein [Homo sapiens])	Contains protein domain (PF00573) - Ribosomal protein L4L1 family	254259, 20281089, 35695052, 265008, 264694, 265011, 264760, 18108351, 264682, 264683, 284389, 264694, 264686, 264687, 264689, 21908796, 264681, 264692, 18108374, 18108377, 264657, 264639, 18108395
2428	85732889 (4853, 4856)	Novel Protein sim. GBank nucleoporin p54 [Rattus norvegicus]		22278989, 22278999, 35898052, 21908794, 264686, 21908795, 21908796, 21908799, 25959391, 25959392, 25959393, 25959394, 25959395, 25959396, 264611, 264609, 21908794, 265018, 265019, 264448, 264683, 264298, 21908795, 21908798, 265021, 264693, 18108381
2429	8769276 (4857, 4858)	Novel Protein sim. GBank neurofilament H [Oryctolagus cuniculus]	Contains protein domain (PF00111) - Beta defensins	UNCLASSIFIED
2430	86348827 (4859, 4860)	Novel Protein sim. GBank gill1537070 (U63840) - nucleoporin p54 [Rattus norvegicus]		UNCLASSIFIED
2431	87649884 (4861, 4862)	Novel Protein sim. GBank gill3860729 (embCAA14630) - (A233270) CELL DIVISION PROTEIN FTSJ (ftsJ) [Rickettsia prowazekii]	Contains protein domain (PF01728) - FtsJ cell division protein	29331826, 29331827, 35696052, 26146499, 264905, 264906, 264901, 264288, 264689, 21908795, 264692, 35696423
2432	80053033 (4853, 4864)	Novel Protein sim. GBank gill3876367 (embCAA33287) - (Z69360) Weak similarity to Emerita fibronectin (FIR Acc. No. A48517); cDNA EST M89266 comes from this gene; cDNA EST M29509.5 comes from this gene [Caenorhabditis elegans]		protease
2433	80055092 (4865, 4866)	Novel Protein sim. GBank gill2224593 (dbjBA20784) - (AF002924) MADA026 [Homo sapiens]	Contains protein domain (PF00596) - Zinc finger, C2H2 type	264469, 264465, 265018, 264762, 264683, 264991, 264595, 264597, 264639, 264595, 264593
2434	19320148 (4867, 4868)	Novel Protein sim. GBank gill1253289 (U47856) - fibrin-4 [for the same species]		UNCLASSIFIED
2435	20730044 (4869, 4870)	Novel Protein sim. GBank gill5941332 (AF091234) - putative transcription factor [Mus musculus]		UNCLASSIFIED
2436	80046068 (4871, 4872)	Novel Protein sim. GBank gill5941332 (AF091234) - putative transcription factor [Mus musculus]		UNCLASSIFIED
2437	83353424 (4873, 4874)	Novel Protein sim. GBank gill5941332 (AF091234) - putative transcription factor [Mus musculus]		UNCLASSIFIED
2438	8413473 (4875, 4876)	Novel Protein sim. GBank gill5941332 (AF091234) - putative transcription factor [Mus musculus]	Contains protein domain (PF01250) - Uncharacterized protein family UPF0031	29331826, 264508, 264905, 264505, 264506, 264506, 264907, 264908, 264909, 264511, 265006, 264512, 264910, 265009, 264591, 33657402, 21906754, 265011, 264760, 264764, 264685, 264686, 264768, 35895917, 33657023, 264693, 264631, 264632, 36182323, 264658, 83373044, 264583, 264584, 264585, 264586, 264587

2439	84850550 (4877, 4878)	Novel Protein sim. GBank g9426515(gpAAU15345) - (ACD04044) small nuclear ribonucleoprotein Sm-D1 (Arabidopsis thaliana)	Contains protein domain (PF01423) - Sm protein	UNCLASSIFIED	60424178, 18100397, 56182575, 56181666, 56994075, 22278996, 35696296, 22278997, 22278999, 264259, 52645080, 29331822, 56182191, 29331824, 60424269, 65714117, 29331825, 60432289, 29331826, 29331827, 29331828, 35696052, 29146498, 264509, 264905, 264906, 52644445, 60431735, 33109954, 11906754, 33657064, 55811366, 52644295, 87166474, 265017, 265018, 265019, 18100351, 264448, 264298, 264788, 52644223, 56181562, 21906764, 21906765, 21906766, 21906767, 21906768, 21906769, 35696917, 33657023, 33657109, 33657182, 21906822, 27403584, 55715135, 264295, 264907, 33109970, 60431538, 264907, 55810754, 35696423, 65274761, 34958455, 60431850, 56182323, 60432113, 22279000, 22279002, 264567
2440	87641733 (4879, 4880)			UNCLASSIFIED	29331822, 29331824, 29331825, 29331827, 35696052, 264506, 264907, 264510, 265018, 265019, 264448, 264369, 265020, 265021, 56182323, 264539, 22279002
2441	87623514 (4881, 4882)	Novel Protein sim. GBank g9302488(gp945241)298_HUMAN - HYPOTHETICAL PROTEIN KIAA0281 (HAB116)		UNCLASSIFIED	264488, 264629, 18100374, 264564
2442	87273590 (4883, 4884)	Novel Protein sim. GBank g9450613(1e1pnc_002703:1)ppp1 - protein phosphatase 1, catalytic subunit	Contains protein domain (PF00560) - Leucine Rich Repeat	UNCLASSIFIED	22278996, 22278999, 29331822, 264768, 264693
2443	84305449 (4885, 4886)	Novel Protein sim. GBank g1170658(gp002975)KID1_RAT - RENAL TRANSCRIPTION FACTOR KID-1 (TRANSCRIPTION FACTOR 17)	Contains protein domain (PF01352) - KRAB box	UNCLASSIFIED	264906
2444	84086345 (4887, 4888)	Novel Protein sim. GBank g94758924(1e1pnc_004280:1)pnrf3 - nuclear factor (erythroid-derived 2)-like 3	Contains protein domain (PF00170) - bZIP transcription factor	UNCLASSIFIED	264259, 18100382, 18100383, 18100385, 22279000
2445	87336536 (4889, 4890)	Novel Protein sim. GBank g12135950(1p1)S58222 - PQ-rich protein - human		UNCLASSIFIED	264259, 35696052, 264569, 18100361
2446	88052933 (4891, 4892)	Novel Protein sim. GBank g94753887(1e1pnc)CA05409.21 - (AJ002424) p65 protein [Homo sapiens]	Contains protein domain (PF00095) - WAP-type (Whey Acidic Protein)	UNCLASSIFIED	265011, 264689, 33657023, 263981, 18100385
2447	94845149 (4893, 4894)	Novel Protein sim. GBank g9485813(1e1pnc_005049:1)stj3 - suppression of tumorigenicity 5	cathecin	UNCLASSIFIED	56182575, 264259, 29331824, 29331825, 29331827, 60433366, 60435438, 264758, 265018, 264692, 65274762, 60431528, 65274791, 56182323

2448	87745860 (4895, 4896)		UNCLASSIFIED	22278986, 22278997, 22278999, 26331826, 35650042, 264107, 264110, 87168474, 87168539, 18108351, 21906767, 21906769, 27486252, 263976, 264459, 264458, 265007, 264459, 265021, 56528486
2449	87769075 (4897, 4898)	Novel Protein sim. GBank g1728837spP39194PALU7_HUMAN - IIII ALU SUBFAMILY SQ VARNING ENTRY IIII	cadherin	
2450	8697764 (4939, 4900)	Novel Protein sim. GBank g1710022spP5528R824_MOUSE - RAS-RELATED PROTEIN RAS-24 (RAB-16)	UNCLASSIFIED	264406, 264093, 26331822, 26331824, 26331825, 6674117, 26331826, 26331828, 35990052, 264907, 66712502, 28331530, 264910, 265008, 264758, 265017, 265016, 264762, 264762, 264763, 264764, 264765, 264766, 33657023, 264693, 336571061, 203959, 83372044, 18108395
2452	91230509 (4903, 4904)	Novel Protein sim. GBank g11504034[bb]BAA13218] - (D65980) KIA0227 [Homo sapiens]	isomerase	264102, 264112, 264688, 263972, 18108374, 83372044, 264563
2453	84201088 (4905, 4906)	Novel Protein sim. GBank g128830079 [AC004142] - similar to murine leucine-rich repeat protein; possible role in neural development by protein-protein interactions; 93% similarity to D45802 (P1D13695805) [Homo sapiens]	Contains protein domain (PF00071) - Leucine Rich Repeat	264509, 264512, 18108385
2454	95101091 (4907, 4908)	Novel Protein sim. GBank g1076802P[349515] - extensin like protein - maize	UNCLASSIFIED	263994, 6674117, 26331827, 264508, 264509, 264905, 264906, 264907, 264908, 264909, 264510, 264511, 264512, 265009, 264910, 264591, 264758, 264759, 265010, 265011, 264603, 264604, 264760, 264761, 264762, 18108351, 264764, 264765, 264766, 264693, 264769, 264534, 264691, 264692, 264693, 264694, 264695, 264696, 264697, 264698, 264699, 264700, 264701, 264702, 264703, 264704, 264705, 264706, 264707, 264708, 264709, 264710, 264711, 264712, 264713, 264714, 264715, 264716, 264717, 264718, 264719, 264720, 264721, 264722, 264723, 264724, 264725, 264726, 264727, 264728, 264729, 264730, 264731, 264732, 264733, 264734, 264735, 264736, 264737, 264738, 264739, 264740, 264741, 264742, 264743, 264744, 264745, 264746, 264747, 264748, 264749, 264750, 264751, 264752, 264753, 264754, 264755, 264756, 264757, 264758, 264759, 264760, 264761, 264762, 264763, 264764, 264765, 264766, 264767, 264768, 264769, 264770, 264771, 264772, 264773, 264774, 264775, 264776, 264777, 264778, 264779, 264780, 264781, 264782, 264783, 264784, 264785, 264786, 264787, 264788, 264789, 264790, 264791, 264792, 264793, 264794, 264795, 264796, 264797, 264798, 264799, 264800, 264801, 264802, 264803, 264804, 264805, 264806, 264807, 264808, 264809, 264810, 264811, 264812, 264813, 264814, 264815, 264816, 264817, 264818, 264819, 264820, 264821, 264822, 264823, 264824, 264825, 264826, 264827, 264828, 264829, 264830, 264831, 264832, 264833, 264834, 264835, 264836, 264837, 264838, 264839, 264840, 264841, 264842, 264843, 264844, 264845, 264846, 264847, 264848, 264849, 264850, 264851, 264852, 264853, 264854, 264855, 264856, 264857, 264858, 264859, 264860, 264861, 264862, 264863, 264864, 264865, 264866, 264867, 264868, 264869, 264870, 264871, 264872, 264873, 264874, 264875, 264876, 264877, 264878, 264879, 264880, 264881, 264882, 264883, 264884, 264885, 264886, 264887, 264888, 264889, 264890, 264891, 264892, 264893, 264894, 264895, 264896, 264897, 264898, 264899, 264900, 264901, 264902, 264903, 264904, 264905, 264906, 264907, 264908, 264909, 264910, 264911, 264912, 264913, 264914, 264915, 264916, 264917, 264918, 264919, 264920, 264921, 264922, 264923, 264924, 264925, 264926, 264927, 264928, 264929, 264930, 264931, 264932, 264933, 264934, 264935, 264936, 264937, 264938, 264939, 264940, 264941, 264942, 264943, 264944, 264945, 264946, 264947, 264948, 264949, 264950, 264951, 264952, 264953, 264954, 264955, 264956, 264957, 264958, 264959, 264960, 264961, 264962, 264963, 264964, 264965, 264966, 264967, 264968, 264969, 264970, 264971, 264972, 264973, 264974, 264975, 264976, 264977, 264978, 264979, 264980, 264981, 264982, 264983, 264984, 264985, 264986, 264987, 264988, 264989, 264990, 264991, 264992, 264993, 264994, 264995, 264996, 264997, 264998, 264999, 265000, 265001, 265002, 265003, 265004, 265005, 265006, 265007, 265008, 265009, 265010, 265011, 265012, 265013, 265014, 265015, 265016, 265017, 265018, 265019, 265020, 265021, 265022, 265023, 265024, 265025, 265026, 265027, 265028, 265029, 265030, 265031, 265032, 265033, 265034, 265035, 265036, 265037, 265038, 265039, 265040, 265041, 265042, 265043, 265044, 265045, 265046, 265047, 265048, 265049, 265050, 265051, 265052, 265053, 265054, 265055, 265056, 265057, 265058, 265059, 265060, 265061, 265062, 265063, 265064, 265065, 265066, 265067, 265068, 265069, 265070, 265071, 265072, 265073, 265074, 265075, 265076, 265077, 265078, 265079, 265080, 265081, 265082, 265083, 265084, 265085, 265086, 265087, 265088, 265089, 265090, 265091, 265092, 265093, 265094, 265095, 265096, 265097, 265098, 265099, 265100, 265101, 265102, 265103, 265104, 265105, 265106, 265107, 265108, 265109, 265110, 265111, 265112, 265113, 265114, 265115, 265116, 265117, 265118, 265119, 265120, 265121, 265122, 265123, 265124, 265125, 265126, 265127, 265128, 265129, 265130, 265131, 265132, 265133, 265134, 265135, 265136, 265137, 265138, 265139, 265140, 265141, 265142, 265143, 265144, 265145, 265146, 265147, 265148, 265149, 265150, 265151, 265152, 265153, 265154, 265155, 265156, 265157, 265158, 265159, 265160, 265161, 265162, 265163, 265164, 265165, 265166, 265167, 265168, 265169, 265170, 265171, 265172, 265173, 265174, 265175, 265176, 265177, 265178, 265179, 265180, 265181, 265182, 265183, 265184, 265185, 265186, 265187, 265188, 265189, 265190, 265191, 265192, 265193, 265194, 265195, 265196, 265197, 265198, 265199, 265200, 265201, 265202, 265203, 265204, 265205, 265206, 265207, 265208, 265209, 265210, 265211, 265212, 265213, 265214, 265215, 265216, 265217, 265218, 265219, 265220, 265221, 265222, 265223, 265224, 265225, 265226, 265227, 265228, 265229, 265230, 265231, 265232, 265233, 265234, 265235, 265236, 265237, 265238, 265239, 265240, 265241, 265242, 265243, 265244, 265245, 265246, 265247, 265248, 265249, 265250, 265251, 265252, 265253, 265254, 265255, 265256, 265257, 265258, 265259, 265260, 265261, 265262, 265263, 265264, 265265, 265266, 265267, 265268, 265269, 265270, 265271, 265272, 265273, 265274, 265275, 265276, 265277, 265278, 265279, 265280, 265281, 265282, 265283, 265284, 265285, 265286, 265287, 265288, 265289, 265290, 265291, 265292, 265293, 265294, 265295, 265296, 265297, 265298, 265299, 265300, 265301, 265302, 265303, 265304, 265305, 265306, 265307, 265308, 265309, 265310, 265311, 265312, 265313, 265314, 265315, 265316, 265317, 265318, 265319, 265320, 265321, 265322, 265323, 265324, 265325, 265326, 265327, 265328, 265329, 265330, 265331, 265332, 265333, 265334, 265335, 265336, 265337, 265338, 265339, 265340, 265341, 265342, 265343, 265344, 265345, 265346, 265347, 265348, 265349, 265350, 265351, 265352, 265353, 265354, 265355, 265356, 265357, 265358, 265359, 265360, 265361, 265362, 265363, 265364, 265365, 265366, 265367, 265368, 265369, 265370, 265371, 265372, 265373, 265374, 265375, 265376, 265377, 265378, 265379, 265380, 265381, 265382, 265383, 265384, 265385, 265386, 265387, 265388, 265389, 265390, 265391, 265392, 265393, 265394, 265395, 265396, 265397, 265398, 265399, 265400, 265401, 265402, 265403, 265404, 265405, 265406, 265407, 265408, 265409, 265410, 265411, 265412, 265413, 265414, 265415, 265416, 265417, 265418, 265419, 265420, 265421, 265422, 265423, 265424, 265425, 265426, 265427, 265428, 265429, 265430, 265431, 265432, 265433, 265434, 265435, 265436, 265437, 265438, 265439, 265440, 265441, 265442, 265443, 265444, 265445, 265446, 265447, 265448, 265449, 265450, 265451, 265452, 265453, 265454, 265455, 265456, 265457, 265458, 265459, 265460, 265461, 265462, 265463, 265464, 265465, 265466, 265467, 265468, 265469, 265470, 265471, 265472, 265473, 265474, 265475, 265476, 265477, 265478, 265479, 265480, 265481, 265482, 265483, 265484, 265485, 265486, 265487, 265488, 265489, 265490, 265491, 265492, 265493, 265494, 265495, 265496, 265497, 265498, 265499, 265500, 265501, 265502, 265503, 265504, 265505, 265506, 265507, 265508, 265509, 265510, 265511, 265512, 265513, 265514, 265515, 265516, 265517, 265518, 265519, 265520, 265521, 265522, 265523, 265524, 265525, 265526, 265527, 265528, 265529, 265530, 265531, 265532, 265533, 265534, 265535, 265536, 265537, 265538, 265539, 265540, 265541, 265542, 265543, 265544, 265545, 265546, 265547, 265548, 265549, 265550, 265551, 265552, 265553, 265554, 265555, 265556, 265557, 265558, 265559, 265560, 265561, 265562, 265563, 265564, 265565, 265566, 265567, 265568, 265569, 265570, 265571, 265572, 265573, 265574, 265575, 265576, 265577, 265578, 265579, 265580, 265581, 265582, 265583, 265584, 265585, 265586, 265587, 265588, 265589, 265590, 265591, 265592, 265593, 265594, 265595, 265596, 265597, 265598, 265599, 265600, 265601, 265602, 265603, 265604, 265605, 265606, 265607, 265608, 265609, 265610, 265611, 265612, 265613, 265614, 265615, 265616, 265617, 265618, 265619, 265620, 265621, 265622, 265623, 265624, 265625, 265626, 265627, 265628, 265629, 265630, 265631, 265632, 265633, 265634, 265635, 265636, 265637, 265638, 265639, 265640, 265641, 265642, 265643, 265644, 265645, 265646, 265647, 265648, 265649, 265650, 265651, 265652, 265653, 265654, 265655, 265656, 265657, 265658, 265659, 265660, 265661, 265662, 265663, 265664, 265665, 265666, 265667, 265668, 265669, 265670, 265671, 265672, 265673, 265674, 265675, 265676, 265677, 265678, 265679, 265680, 265681, 265682, 265683, 265684, 265685, 265686, 265687, 265688, 265689, 265690, 265691, 265692, 265693, 265694, 265695, 265696, 265697, 265698, 265699, 265700, 265701, 265702, 265703, 265704, 265705, 265706, 265707, 265708, 265709, 265710, 265711, 265712, 265713, 265714, 265715, 265716, 265717, 265718, 265719, 265720, 265721, 265722, 265723, 265724, 265725, 265726, 265727, 265728, 265729, 265730, 265731, 265732, 265733, 265734, 265735, 265736, 265737, 265738, 265739, 265740, 265741, 265742, 265743, 265744, 265745, 265746, 265747, 265748, 265749, 265750, 265751, 265752, 265753, 265754, 265755, 265756, 265757, 265758, 265759, 265760, 265761, 265762, 265763, 265764, 265765, 265766, 265767, 265768, 265769, 265770, 265771, 265772, 265773, 265774, 265775, 265776, 265777, 265778, 265779, 265780, 265781, 265782, 265783, 265784, 265785, 265786, 265787, 265788, 265789, 265790, 265791, 265792, 265793, 265794, 265795, 265796, 265797, 265798, 265799, 265800, 265801, 265802, 265803, 265804, 265805, 265806, 265807, 265808, 265809, 265810, 265811, 265812, 265813, 265814, 265815, 265816, 265817, 265818, 265819, 265820, 265821, 265822, 265823, 265824, 265825, 265826, 265827, 265828, 265829, 265830, 265831, 265832, 265833, 265834, 265835, 265836, 265837, 265838, 265839, 265840, 265841, 265842, 265843, 265844, 265845, 265846, 265847, 265848, 265849, 265850, 265851, 265852, 265853, 265854, 265855, 265856, 265857, 265858, 265859, 265860, 265861, 265862, 265863, 265864, 265865, 265866, 265867, 265868, 265869, 265870, 265871, 265872, 265873, 265874, 265875, 265876, 265877, 265878, 265879, 265880, 265881, 265882, 265883, 265884, 265885, 265886, 265887, 265888, 265889, 265890, 265891, 265892, 265893, 265894, 265895, 265896, 265897, 265898, 265899, 265900, 265901, 265902, 265903, 265904, 265905, 265906, 265907, 265908, 265909, 265910, 265911, 265912, 265913, 265914, 265915, 265916, 265917, 265918, 265919, 265920, 265921, 265922, 265923, 265924, 265925, 265926, 265927, 265928, 265929, 265930, 265931, 265932, 265933, 265934, 265935, 265936, 265937, 265938, 265939, 265940, 265941, 265942, 265943, 265944, 265945, 265946, 265947, 265948, 265949, 265950, 265951, 265952, 265953, 265954, 265955, 265956, 265957, 265958, 265959, 265960, 265961, 265962, 265963, 265964, 265965, 265966, 265967, 265968, 265969, 265970, 265971, 265972, 265973, 265974, 265975, 265976, 265977, 265978, 265979, 265980, 265981, 265982, 265983, 265984, 265985, 265986, 265987, 265988, 265989, 265990, 265991, 265992, 265993, 265994, 265995, 265996, 265997, 265998, 265999, 266000, 266001, 266002, 266003, 266004, 266005, 266006, 266007, 266008, 266009, 266010, 266011, 266012, 266013, 266014, 266015, 266016, 266017, 266018, 266019, 266020, 266021, 266022, 266023, 266024, 266025, 266026, 266027, 266028, 266029, 266030, 266031, 266032, 266033, 266034, 266035, 266036, 266037, 266038, 266039, 266040, 266041, 266042, 266043, 266044, 266045, 266046, 266047, 266048, 266049, 266050,

2457	94118375 (4913, 4914)	Novel Protein sim. GBank gij205447 (AC004326) - [532184_3] [Homo sapiens]	UNCLASSIFIED	56181686, 264905, 264907, 264511, 264586, 58811386, 264682, 264684, 264685, 264687, 264691, 33657023, 264693, 35695855, 264636, 264555, 56182323, 264558, 56506486, 264503
2458	55515304 (4915, 4916)	Novel Protein sim. GBank gij2364942 (AF022885) - Similar to collagen [Caenorhabditis elegans]	UNCLASSIFIED	264691, 264693, 264694, 264559
2459	87551915 (4917, 4918)	Novel Protein sim. GBank gij54419429 (AF004987) - Supported by mouse EST AA530043 (ND g2284036) [Homo sapiens]	UNCLASSIFIED	2746265
2460	94315289 (4919, 4920)	Novel Protein sim. GBank gij49287019 (AF020411) - [AF151874] CQ1-16 protein [Homo sapiens]	UNCLASSIFIED	65274572, 35690285, 22278696, 22278697, 60432049, 29331822, 29331824, 29331825, 6674117, 60432289, 29331826, 29331827, 29331828, 33655970, 29164699, 264102, 294109, 60433438, 265017, 265018, 265019, 264288, 21906765, 21906766, 21906769, 35695917, 265020, 264691, 33657023, 27488261, 18108374, 35695855, 87168518, 60432113
2461	87451547 (4921, 4922)	Novel Protein sim. GBank gij44269229 (AF120662) - [AF120662] Arf-like 2 binding protein BART1 [Homo sapiens]	UNCLASSIFIED	264259, 29331828, 264910, 18108351, 18108370, 18108374
2462	86568302 (4923, 4924)	Novel Protein sim. GBank gij5425387 (AF145613) - [AF145613] [Homo sapiens]	UNCLASSIFIED	264909, 264758, 264684, 18108374, 264637, 18108385
2463	84388543 (4925, 4926)	Novel Protein sim. GBank gij509251 (AF145613) - [AF145613] [Homo sapiens]	UNCLASSIFIED	264681, 264566
2464	91719957 (4927, 4928)	Novel Protein sim. GBank gij54103009 (AF100757) - [AF100757] COP9 complex subunit 4 [Homo sapiens]	UNCLASSIFIED	264480, 52646842, 22278995, 35695298, 22278996, 22278997, 22278999, 264259, 29331822, 29331824, 6674117, 29331825, 29331826, 29331827, 35696052, 29331828, 264509, 52644045, 264510, 264511, 264512, 265008, 60170831, 264593, 52646317, 33109954, 33857084, 265017, 265018, 265019, 264762, 264448, 264764, 264688, 264766, 21906765, 21906766, 21906767, 21906768, 21906769, 265021, 33657023, 33657109, 18108370, 18108381, 60170394, 18108385, 22279602, 264488

2465	55357483 (4928, 4930)	Novel Protein sim. GBank gH506-401[pepNP_002871]-v-raf-1 murine leukemia viral oncogene homolog 1	oncogene	Contains protein domain (PF00068) Eukaryotic protein kinase domain	22103392, 52645150, 52645156, 55845365, 55845369, 55845370, 55845371, 55845372, 55845373, 55845374, 55845375, 55845376, 55845377, 55845378, 55845379, 55845380, 55845381, 55845382, 55845383, 55845384, 55845385, 55845386, 55845387, 55845388, 55845389, 55845390, 55845391, 55845392, 55845393, 55845394, 55845395, 55845396, 55845397, 55845398, 55845399, 55845400, 55845401, 55845402, 55845403, 55845404, 55845405, 55845406, 55845407, 55845408, 55845409, 55845410, 55845411, 55845412, 55845413, 55845414, 55845415, 55845416, 55845417, 55845418, 55845419, 55845420, 55845421, 55845422, 55845423, 55845424, 55845425, 55845426, 55845427, 55845428, 55845429, 55845430, 55845431, 55845432, 55845433, 55845434, 55845435, 55845436, 55845437, 55845438, 55845439, 55845440, 55845441, 55845442, 55845443, 55845444, 55845445, 55845446, 55845447, 55845448, 55845449, 55845450, 55845451, 55845452, 55845453, 55845454, 55845455, 55845456, 55845457, 55845458, 55845459, 55845460, 55845461, 55845462, 55845463, 55845464, 55845465, 55845466, 55845467, 55845468, 55845469, 55845470, 55845471, 55845472, 55845473, 55845474, 55845475, 55845476, 55845477, 55845478, 55845479, 55845480, 55845481, 55845482, 55845483, 55845484, 55845485, 55845486, 55845487, 55845488, 55845489, 55845490, 55845491, 55845492, 55845493, 55845494, 55845495, 55845496, 55845497, 55845498, 55845499, 55845500, 55845501, 55845502, 55845503, 55845504, 55845505, 55845506, 55845507, 55845508, 55845509, 55845510, 55845511, 55845512, 55845513, 55845514, 55845515, 55845516, 55845517, 55845518, 55845519, 55845520, 55845521, 55845522, 55845523, 55845524, 55845525, 55845526, 55845527, 55845528, 55845529, 55845530, 55845531, 55845532, 55845533, 55845534, 55845535, 55845536, 55845537, 55845538, 55845539, 55845540, 55845541, 55845542, 55845543, 55845544, 55845545, 55845546, 55845547, 55845548, 55845549, 55845550, 55845551, 55845552, 55845553, 55845554, 55845555, 55845556, 55845557, 55845558, 55845559, 55845560, 55845561, 55845562, 55845563, 55845564, 55845565, 55845566, 55845567, 55845568, 55845569, 55845570, 55845571, 55845572, 55845573, 55845574, 55845575, 55845576, 55845577, 55845578, 55845579, 55845580, 55845581, 55845582, 55845583, 55845584, 55845585, 55845586, 55845587, 55845588, 55845589, 55845590, 55845591, 55845592, 55845593, 55845594, 55845595, 55845596, 55845597, 55845598, 55845599, 55845600, 55845601, 55845602, 55845603, 55845604, 55845605, 55845606, 55845607, 55845608, 55845609, 55845610, 55845611, 55845612, 55845613, 55845614, 55845615, 55845616, 55845617, 55845618, 55845619, 55845620, 55845621, 55845622, 55845623, 55845624, 55845625, 55845626, 55845627, 55845628, 55845629, 55845630, 55845631, 55845632, 55845633, 55845634, 55845635, 55845636, 55845637, 55845638, 55845639, 55845640, 55845641, 55845642, 55845643, 55845644, 55845645, 55845646, 55845647, 55845648, 55845649, 55845650, 55845651, 55845652, 55845653, 55845654, 55845655, 55845656, 55845657, 55845658, 55845659, 55845660, 55845661, 55845662, 55845663, 55845664, 55845665, 55845666, 55845667, 55845668, 55845669, 55845670, 55845671, 55845672, 55845673, 55845674, 55845675, 55845676, 55845677, 55845678, 55845679, 55845680, 55845681, 55845682, 55845683, 55845684, 55845685, 55845686, 55845687, 55845688, 55845689, 55845690, 55845691, 55845692, 55845693, 55845694, 55845695, 55845696, 55845697, 55845698, 55845699, 55845700, 55845701, 55845702, 55845703, 55845704, 55845705, 55845706, 55845707, 55845708, 55845709, 55845710, 55845711, 55845712, 55845713, 55845714, 55845715, 55845716, 55845717, 55845718, 55845719, 55845720, 55845721, 55845722, 55845723, 55845724, 55845725, 55845726, 55845727, 55845728, 55845729, 55845730, 55845731, 55845732, 55845733, 55845734, 55845735, 55845736, 55845737, 55845738, 55845739, 55845740, 55845741, 55845742, 55845743, 55845744, 55845745, 55845746, 55845747, 55845748,
------	-----------------------	---	----------	---	--

2475	94321693 (4949, 4950)	Novel Protein sim. G.Bank gll1216486 (J48452) - HT protein (Cristallus griseus)	Contains protein domain (PF00038) - Ig EGF-like domain	264259, 29331822, 265006, 265007, 265010, 265011, 264448, 264286, 264368, 264685, 264686, 18108357, 264768, 18108362, 264693, 18108370, 18108374, 18108379, 34966423, 83373044, 18108383, 18108385, 264564, 264565, 264567
2476	94315618 (4951, 4952)	Novel Protein sim. G.Bank gll252827 (AC004382) - (Unknown gene product [Homo sapiens])	UNCLASSIFIED	264259, 60422269, 2671417, 264905, 265006, 264511, 265008, 265009, 264756, 265010, 265011, 18108351, 264681, 264389, 264285, 264689, 21906707, 265020, 18108374, 264633, 18108382, 83373044, 18108385, 87169316
2477	20718974 (4953, 4954)		UNCLASSIFIED	263976
2478	17659165 (4955, 4956)		UNCLASSIFIED	265017
2479	94314569 (4967, 4968)	Novel Protein sim. G.Bank gll16442232 [BAX11032] - (D57066) N-WASP [Box. laurus]	UNCLASSIFIED	56904075, 22278998, 21906754, 264632, 21906765
2480	95295605 (4958, 4960)	Novel Protein sim. G.Bank gll168946 [gll168946A33018.1] - (AB028989) KIAA1066 protein [Homo sapiens]	UNCLASSIFIED	264805, 264807, 264765
2481	94718481 (4961, 4962)		collagen	63274572, 56182575, 22278997, 264094, 264259, 29331822, 29331824, 66714117, 29331827, 35696052, 264508, 264905, 264906, 264807, 264908, 52644045, 264909, 56182435, 265008, 264910, 33657402, 55812038, 264758, 265010, 265011, 265017, 265018, 264760, 267162, 18108351, 264764, 264285, 264766, 264686, 264763, 21906768, 55811957, 265020, 264681, 264682, 264683, 264628, 55811576, 264630, 264634, 264635, 264636, 264637, 264638, 264639, 83373044, 60432113, 22278992, 83373044, 60432113, 22278992
2482	87393165 (4963, 4964)	Novel Protein sim. G.Bank gll121249 [gll121249] - guanine nucleotide-exchange activator CDC25 homolog - mouse	UNCLASSIFIED	29331832, 29331824, 29331825, 29331827, 264508, 264905, 264509, 264906, 264807, 264908, 264611, 264591, 264763, 264693, 264631, 264632, 264636, 264638, 264639, 264563
2483	87731833 (4965, 4966)		UNCLASSIFIED	264485, 22278995, 264093, 264995, 60432049, 60433356, 60433438, 264448, 264286, 263967, 18108370, 18108385, 18108388, 264482
2484	94187774 (4967, 4968)	Novel Protein sim. G.Bank gll128831 [gll128831] - HUMAN - III ALU SUBFAMILY JWARNING ENTRY III	kinase	264563
2485	87166556 (4969, 4970)	Novel Protein sim. G.Bank gll163397 (J25281) - SH3 domain binding protein [Rattus norvegicus]	UNCLASSIFIED	22278995, 22278996, 22278997, 22278999, 264539, 60432049, 29331824, 60432049, 264539, 264540, 264603, 264604, 264510, 264444, 264486, 264685, 264686, 264769, 264689, 35655917, 265022, 264692, 264693, 56182323

2483	35422415 (4985, 4986)	Novel Protein sim. GBank g14203070[bi]BAX7432.1] - (AB020716) KIAA0909 protein [Homo sapiens]	Contains protein domain (PF01424) - R3H domain	18108394, 264687, 55774572, 56182575, 22278995, 56994075, 60420249, 29331822, 29331824, 29331825, 29331826, 29331827, 2914648, 264508, 264509, 264510, 264505, 264807, 29331830, 264509, 264510, 265005, 264511, 265007, 264512, 265008, 265009, 264910, 21906754, 265011, 264600, 265017, 265018, 264604, 264605, 265019, 5581150, 264782, 18108351, 264681, 264448, 264683, 264389, 264288, 18108355, 18108357, 264687, 21906765, 21906766, 21906767, 21906768, 21906769, 265520, 264691, 264692, 33657023, 33657349, 18108370, 18108374, 18108376, 336570794, 18108379, 18108380, 264555, 264556, 264557, 264558, 264636, 264555, 264637, 264557, 264558, 264639, 264559, 83373044, 18108385, 87188518, 60432113, 22279600, 22279602, 264482, 264556, 264485
2484	30793118 (4887, 4888)	Novel Protein sim. GBank g14203070[bi]BAX7432.1] - (AJ243460) proteoglycan [Leishmania major]	UNCLASSIFIED	264807, 264561
2485	94234551 (4889, 4990)	Novel Protein sim. GBank g14203070[bi]BAX7432.1] - (AJ243460) proteoglycan [Leishmania major]	collagen	263884, 22278997, 33656552, 264509, 264505, 264906, 264907, 264908, 264909, 265006, 265009, 264595, 264604, 264448, 264682, 264794, 264288, 264685, 264768, 264789, 264689, 265020, 264692, 63274620, 264629, 55810794, 33696423, 55811576, 264636, 264637, 18108385, 22279600, 264584, 264587, 264485
2486	80018765 (4891, 4892)	Novel Protein sim. GBank g14203070[bi]BAX7432.1] - (AJ223116) Q11773.1 (PUTATIVE novel protein) [Homo sapiens]	slud	2814760, 264905, 265005, 265007, 18108384, 18108382, 18108370, 18108374, 264555, 264556, 18108391, 18108393, 18108398
2487	91723554 (4893, 4994)	Novel Protein sim. GBank g14203070[bi]BAX7432.1] - (AJ223116) Q11773.1 (PUTATIVE novel protein) [Homo sapiens]	UNCLASSIFIED	52844507, 22278998, 22278999, 26531824, 26531828, 33657402, 21906754, 87168474, 265019, 264389, 264689, 21906785, 21906786, 21906767, 21906768, 265020, 33657023, 18108376, 18108387
2488	87724633 (4895, 4995)	Novel Protein sim. GBank g11200503 (U47824) - B [Homo sapiens]	UNCLASSIFIED	29331827, 264512, 264910, 264428, 18108374, 35958595
2489	94655125 (4997, 4998)	Novel Protein sim. GBank g13510234 (AC005581) - [R37237.1, partial CDS [Homo sapiens]	Contains protein domain (PF00089) - Eukaryotic protein kinase domain	264909, 53812038, 264631, 264637, 264558

2500	94643324 (4999, 5000)	Novel Protein sim. GBank gi 3881275 em CAA31725 - (A1032855) predicted using Genefinder, similar to Inositol monophosphatase family, cDNA EST X2256:11.5 comes from this gene [Caenorhabditis elegans]	Contains protein domain (PF004159) - Inositol monophosphatase family	transport	52644507, 52645156, 22278995, 56904075, 35692386, 22278998, 264295, 52645090, 29331824, 29331825, 69714117, 60432289, 29331826, 29331827, 35690552, 29331828, 264508, 264500, 264510, 264512, 33657402, 60433433, 21900754, 52644256, 87168474, 87163559, 264603, 264681, 264448, 264683, 264286, 264369, 52644229, 264639, 21900765, 21900766, 21900767, 21900768, 21900769, 55811557, 35695917, 265020, 265021, 52644150, 33657023, 264683, 33657182, 35695763, 35696423, 35695855, 52644332, 83373044, 18108387, 87168518, 22279002, 55811573, 35696036, 22278996, 35696037, 22278997, 29331827, 29331828, 29331829, 29331830, 29331831, 29331832, 29331833, 29331834, 29331835, 29331836, 29331837, 35696042, 264905, 264906, 264907, 264908, 264909, 56182435, 264510, 264511, 265007, 264910, 264591, 60432229, 33657402, 60433356, 264595, 55612038, 264758, 264596, 87168474, 87168559, 264600, 264601, 264602, 265017, 264604, 265018, 264605, 265019, 18108351, 264448, 264369, 264288, 264766, 18108357, 21900785, 21900786, 21900787, 21900789, 20148629, 35695917, 264692, 33657023, 264629, 35696423, 55811576, 35695955, 264630, 264634, 264635, 264636, 264638, 264639, 264640, 6070394, 83373044, 18108355, 18108367, 87168516, 60432113, 22279002, 18108368, 264697, 56182435, 265007, 55274572, 264907, 56182435, 265007, 264592, 264760, 18108351, 264448, 264369, 264288, 264684, 264686, 55811557, 265021, 264692, 33657109, 2933973, 55811576, 264635, 264555, 264556, 264557, 264558, 56182323, 264559, 87168518, 264583, 264482, 264805, 264807, 264512, 265008, 265011, 18108351, 264448, 264288, 29148627, 264693, 18108370, 18108374, 18108385
2501	94303896 (5001, 5002)	Novel Protein sim. GBank gi 4929615 gp AD34068 (AF151833 - (AF151831) CGI-73 protein [Homo sapiens])	Contains protein domain (PF00651) - BTB/POZ domain		55811573, 35696036, 22278996, 35696037, 22278997, 29331827, 29331828, 29331829, 29331830, 29331831, 29331832, 29331833, 29331834, 29331835, 29331836, 29331837, 35696042, 264905, 264906, 264907, 264908, 264909, 56182435, 264510, 264511, 265007, 264910, 264591, 60432229, 33657402, 60433356, 264595, 55612038, 264758, 264596, 87168474, 87168559, 264600, 264601, 264602, 265017, 264604, 265018, 264605, 265019, 18108351, 264448, 264369, 264288, 264766, 18108357, 21900785, 21900786, 21900787, 21900789, 20148629, 35695917, 264692, 33657023, 264629, 35696423, 55811576, 35695955, 264630, 264634, 264635, 264636, 264638, 264639, 264640, 6070394, 83373044, 18108355, 18108367, 87168516, 60432113, 22279002, 18108368, 264697, 56182435, 265007, 55274572, 264907, 56182435, 265007, 264592, 264760, 18108351, 264448, 264369, 264288, 264684, 264686, 55811557, 265021, 264692, 33657109, 2933973, 55811576, 264635, 264555, 264556, 264557, 264558, 56182323, 264559, 87168518, 264583, 264482, 264805, 264807, 264512, 265008, 265011, 18108351, 264448, 264288, 29148627, 264693, 18108370, 18108374, 18108385
2502	99993716 (5003, 5004)	Novel Protein sim. GBank gi 3041847 AC004542 - OXYSTEROL-BINDING PROTEIN-like; similar to P22059 (P1Dg29308) [Homo sapiens]	Contains protein domain (PF01237) - UNCLASSIFIED Oysterol-binding protein		55811573, 35696036, 22278996, 35696037, 22278997, 29331827, 29331828, 29331829, 29331830, 29331831, 29331832, 29331833, 29331834, 29331835, 29331836, 29331837, 35696042, 264905, 264906, 264907, 264908, 264909, 56182435, 264510, 264511, 265007, 264910, 264591, 60432229, 33657402, 60433356, 264595, 55612038, 264758, 264596, 87168474, 87168559, 264600, 264601, 264602, 265017, 264604, 265018, 264605, 265019, 18108351, 264448, 264369, 264288, 264766, 18108357, 21900785, 21900786, 21900787, 21900789, 20148629, 35695917, 264692, 33657023, 264629, 35696423, 55811576, 35695955, 264630, 264634, 264635, 264636, 264638, 264639, 264640, 6070394, 83373044, 18108355, 18108367, 87168516, 60432113, 22279002, 18108368, 264697, 56182435, 265007, 55274572, 264907, 56182435, 265007, 264592, 264760, 18108351, 264448, 264369, 264288, 264684, 264686, 55811557, 265021, 264692, 33657109, 2933973, 55811576, 264635, 264555, 264556, 264557, 264558, 56182323, 264559, 87168518, 264583, 264482, 264805, 264807, 264512, 265008, 265011, 18108351, 264448, 264288, 29148627, 264693, 18108370, 18108374, 18108385
2503	87878345 (5005, 5006)	Novel Protein sim. GBank gi 2198874 ems CAJ2638 - (N11898) BRX protein [Mus musculus]			55811573, 35696036, 22278996, 35696037, 22278997, 29331827, 29331828, 29331829, 29331830, 29331831, 29331832, 29331833, 29331834, 29331835, 29331836, 29331837, 35696042, 264905, 264906, 264907, 264908, 264909, 56182435, 264510, 264511, 265007, 264910, 264591, 60432229, 33657402, 60433356, 264595, 55612038, 264758, 264596, 87168474, 87168559, 264600, 264601, 264602, 265017, 264604, 265018, 264605, 265019, 18108351, 264448, 264369, 264288, 264766, 18108357, 21900785, 21900786, 21900787, 21900789, 20148629, 35695917, 264692, 33657023, 264629, 35696423, 55811576, 35695955, 264630, 264634, 264635, 264636, 264638, 264639, 264640, 6070394, 83373044, 18108355, 18108367, 87168516, 60432113, 22279002, 18108368, 264697, 56182435, 265007, 55274572, 264907, 56182435, 265007, 264592, 264760, 18108351, 264448, 264369, 264288, 264684, 264686, 55811557, 265021, 264692, 33657109, 2933973, 55811576, 264635, 264555, 264556, 264557, 264558, 56182323, 264559, 87168518, 264583, 264482, 264805, 264807, 264512, 265008, 265011, 18108351, 264448, 264288, 29148627, 264693, 18108370, 18108374, 18108385

2504	87668706 (5007, 5008)	Novel Protein sim. GBank gi1550420(emc)(CAA46220) - (X68107) (ng) (Rattus norvegicus)		264488, 52644507, 52645156, 52646842, 22278094, 2634259, 52645080, 29331822, 29331824, 29331825, 29331826, 29331827, 35690052, 264906, 264908, 52644045, 265009, 60433356, 33657402, 60433438, 264595, 33109854, 87168474, 265017, 265019, 264448, 264428, 264766, 52644229, 21900765, 21906766, 21906767, 21906768, 52644150, 264692, 27468281, 27468282, 27468284, 27468285, 35695763, 35696423, 35695855, 52644332, 56182323, 18108387, 87168518, 60432113, 22279002, 264564
2505	8760559 (5009, 5010)			264488, 52644507, 52645156, 52646842, 22278094, 2634259, 52645080, 29331822, 29331824, 29331825, 29331826, 29331827, 35690052, 264906, 264908, 52644045, 265009, 60433356, 33657402, 60433438, 264595, 33109854, 87168474, 265017, 265019, 264448, 264428, 264766, 52644229, 21900765, 21906766, 21906767, 21906768, 52644150, 264692, 27468281, 27468282, 27468284, 27468285, 35695763, 35696423, 35695855, 52644332, 56182323, 18108387, 87168518, 60432113, 22279002, 264564
2506	91232326 (5011, 5012)	Novel Protein sim. GBank gi1237562(gp)(J49635 - mouse Dhmt protein - mouse		UNCLASSIFIED nuclease
2507	95316233 (5013, 5014)	Novel Protein sim. GBank gi1574489(emc)_006035.1 (pKIAA - histone deacetylase 6		264488, 52644507, 52645156, 52646842, 22278094, 2634259, 52645080, 29331822, 29331824, 29331825, 29331826, 29331827, 35690052, 264906, 264908, 52644045, 265009, 60433356, 33657402, 60433438, 264595, 33109854, 87168474, 265017, 265019, 264448, 264428, 264766, 52644229, 21900765, 21906766, 21906767, 21906768, 52644150, 264692, 27468281, 27468282, 27468284, 27468285, 35695763, 35696423, 35695855, 52644332, 56182323, 18108387, 87168518, 60432113, 22279002, 264564
2508	95316566 (5015, 5016)	Novel Protein sim. GBank gi142643(emc)(CA942809 1) - (AL031447) dJ126A3.2.1 (novel protein) (isolom 1) (homo sapiens)		UNCLASSIFIED
2509	87613741 (5017, 5018)	Novel Protein sim. GBank gi1263289 (U47856) - (Raneus diadematus)		UNCLASSIFIED

2510	95421370 (5018, 5020)	Novel Protein sim. GBank gll325357jlpAC25762.1 - (AF071059) zinc finger RNA binding protein [Mus musculus]	dna_ma_bind	6374572, 22778994, 22278996, 22278998, 22278999, 60432048, 264459, 26331822, 26331824, 26331825, 6671417, 60432289, 26331836, 33656970, 264505, 66712592, 265007, 264810, 60770831, 60432229, 60433356, 60433438, 21906754, 8716474, 265017, 265018, 264448, 264288, 21906767, 21906768, 21906769, 55811957, 35693917, 265020, 265022, 264691, 33657023, 264693, 63274620, 33657100, 33657182, 27486262, 33657349, 18100370, 33659595, 264595, 5818233, 6337304, 6042113, 22279002, 265017, 21906754, 263022, 264692
2511	87384281 (5021, 5022)	Novel Protein sim. GBank gll423252jlpAC4015228.1 - (AF068863) Ets protein, Src C [Mus musculus]		
2512	86084771 (5023, 5024)	Novel Protein sim. GBank gll4502075jlpAC001133 - 1pAMFR - autocrine motility factor receptor	Contains protein domain (PF00097) - transport Zinc finger, C3H-C4 type (RING finger)	22278999, 264259, 26331825, 26331826, 26446499, 264907, 264909, 265008, 265008, 264591, 60432229, 21906754, 26331827, 264693, 264766, 18100357, 264693, 21906769, 264693, 18100370, 263972, 18100374, 264558, 22279000
2513	95357843 (5025, 5026)	Novel Protein sim. GBank gll30046857 (AF071777) - boboy sox [Drosophila melanogaster]	UNCLASSIFIED	60424179, 52645156, 18100394, 22278994, 33696286, 56994075, 22278996, 29331822, 29331824, 60424269, 29331825, 29331827, 33656970, 60431735, 33657084, 87164559, 265017, 264448, 264369, 56181962, 21906766, 21906767, 21906768, 21906769, 265020, 265021, 33657023, 18100396, 33657109, 27486261, 27486262, 33657349, 18100374, 55810764, 35696423, 56182323, 33657350, 18100355
2514	88394575 (5027, 5028)	Novel Protein sim. GBank gll258437 (AF008197) - syncollin [Rattus norvegicus]	UNCLASSIFIED	264510
2515	87394503 (5028, 5030)	Novel Protein sim. GBank gll3575772jlpAC181833 - (AL022727) dJ0018.7 (olfactory receptor-like protein (hSMT-3)) [Homo sapiens]	Contains protein domain (PF00001) - htr7 7 transmembrane receptor (rhodopsin family)	264259, 29146486, 264305, 264786, 29146529, 35695991, 27486261, 264634, 264691, 29331824, 33631825, 29331836, 29331838, 33696053, 264508, 264509, 264505, 264906, 264907, 264908, 264909, 264511, 264910, 33657402, 264757, 33109994, 265017, 265018, 256405, 264760, 264782, 264763, 264766, 264768, 264769, 33657109, 33657182, 264628, 55811570, 35696423, 264631, 264634, 264637, 264638, 264639, 87169518, 22278902, 264594
2516	87766900 (5031, 5032)	Novel Protein sim. GBank gll220552jlpAC230001 - (AL035356) putative protein [Arabidopsis thaliana]	UNCLASSIFIED	
2517	87784960 (5033, 5034)	Novel Protein sim. GBank gll220552jlpAC230001 - (AL035356) putative protein [Arabidopsis thaliana]	UNCLASSIFIED	

2518	9417410 (5035, 5036)	Novel Protein sim. GBank gi4925591 g AACDQ058.1 AF15181 - (AF15181) CGI-61 protein [Homo sapiens]	Contains protein domain (PF00018) - SH3 domain	UNCLASSIFIED	35696266, 264259, 29331822, 29331824, 29331826, 29331829, 29331826, 29331827, 35696055, 29331828, 264907, 264909, 264511, 265007, 60432229, 60433356, 60433438, 55812038, 265010, 265017, 264448, 264286, 264689, 21906768, 21900769, 265022, 52644150, 264693, 18108370, 263972, 264595, 56182323, 93373044, 16108345, 60432113, 264089 284293, 60714117, 29331828, 29331827, 29331826, 264907, 60712502, 265018, 264909, 264907, 60712502, 265018, 55811576, 65274701, 264632, 264555, 264536, 22273902, 264564
2519	94329180 (5037, 5038)	Novel Protein sim. GBank (AC004483) similar to KIAA0766, similar to PIDG382233 [Homo sapiens]	kinase		294259, 264908, 264910, 264482, 21906769, 265020, 264583
2520	87413235 (5039, 5040)	Novel Protein sim. GBank gi4526722 g NP_005685.1 PFATP - fatty acid transport protein 4	transport		
2521	95316244 (5041, 5042)	Novel Protein sim. GBank gi5174488 g NP_005035.1 PKIAA - histone deacetylase 8	Contains protein domain (PF00850) - Histone deacetylase family	Histone	264488, 264489, 263994, 65274572, 22278955, 22278996, 264259, 29331822, 29331826, 264508, 264405, 264509, 264906, 264907, 66712502, 264511, 265006, 265007, 264591, 264592, 264593, 264594, 264595, 264596, 264681, 264440, 264763, 264682, 264764, 264684, 264393, 264288, 264605, 254696, 21900768, 55811957, 264692, 254693, 21908261, 15108370, 264693, 264693, 16108374, 55811957, 55811953, 55811952, 55811951, 1038365, 65274727, 60432113, 264563, 264564, 264585, 264596, 264567
2522	87145652 (5043, 5044)	Novel Protein sim. GBank gi4549001 g NP_004291.1 ILU (02 - (U81002) TRAF4 associated factor 1 [Homo sapiens]	transport		264489, 22278997, 20281171, 21900754, 35695917, 263987, 263976, 263981, 20281169
2523	95340467 (5045, 5046)	Novel Protein sim. GBank gi1805327 g U76374 - skin- BOF2 [Mus musculus]	Contains protein domain (PF01753) - MYND finger		263989
2524	85340469 (5047, 5048)	Novel Protein sim. GBank gi1805327 g U76374 - skin- BOF2 [Mus musculus]			56994075, 22278996, 35662688, 22278997, 22278998, 22278999, 264259, 29331822, 29331824, 29331826, 29331828, 29331830, 55182435, 264512, 265006, 60710831, 33657402, 265010, 87168559, 265019, 264288, 21900765, 21906769, 35695917, 265020, 265021, 265022, 52644150, 264691, 33657023, 33697108, 27466261, 35695423, 65274791, 264595, 53373044, 56264606, 87168518, 264567

2534	81332322 (5067, 5068)	Novel Protein sim. GBank g1042473 (AF094209) - serine/threonine protein kinase TAO1 (Rattus norvegicus)	UNCLASSIFIED	264255, 35696052, 264905, 355017, 21906769, 265020, 265022, 35657109, 22279000
2535	91225058 (5069, 5070)	Novel Protein sim. GBank g1446831 (jmbh/CAB37892) - (AL031430) L4US5X24.2.1 (PUTATIVE novel protein) (isoform 1) [Homo sapiens]		65274572, 35696286, 30432285, 29331828, 265018, 265019, 264288, 264689, 264689, 21908798, 265020, 265021, 264656, 60170394, 22279002
2538	94218540 (5071, 5072)	Novel Protein sim. GBank g17288363 (JP39193) ALU6 - III ALU SUBFAMILY SP WARNING ENTRY III	Mirasep	18108351, 264681, 264683, 264685, 264686, 264687, 264688, 264689, 264690, 264691, 264696, 50246317, 18108351, 264681, 264683, 18108354, 264288, 264687, 264769, 264689, 21906765, 21906766, 21906767, 255021, 50246129, 33657109, 18108374, 18108380, 56182323, 18108381, 18108386, 87168516, 60432113, 22279000, 22279002, 264697, 18108391
2537	95422283 (5073, 5074)	Novel Protein sim. GBank g1455102 (epjnk_000913_1) pHEPC - guanine nucleotide exchange factor p52	ubiquitin	65274572, 35696286, 29331822, 29331825, 29331827, 29331828, 35696052, 264905, 66712502, 264909, 265008, 265011, 264760, 264288, 284885, 35695917, 60170015, 264691, 35657023, 65274620, 35657109, 18108374, 35696423, 35696855, 264636, 264638, 60170394, 26132223, 63373044
2538	36853464 (5075, 5076)		UNCLASSIFIED	22278996, 22278999, 29331822, 29331825, 29331828, 29145496, 264608, 264112, 60170031, 87168559, 264604, 265019, 264685, 264766, 87168516, 22279000, 264655, 264566
2539	94144916 (5077, 5078)		UNCLASSIFIED	22278997, 29331823, 265008, 265009, 284758, 265010, 18108351, 264683, 264288, 21906765, 35695917, 265020, 18108374, 264697
2540	94218545 (5079, 5080)	Novel Protein sim. GBank g11362647 (jplj/S53876 - sex-pseudobaculura)	UNCLASSIFIED	264488, 264768, 264689, 264511, 20281171, 294634, 264635, 264691, 264635, 29331824, 264603, 264604, 264905, 264907, 264608, 264768
2541	95308230 (5081, 5082)	Novel Protein sim. GBank g1171865 (jplj/71865) MOUSE - SEX/THR RICH PROTEIN T10 IN DCCR REGION	UNCLASSIFIED	

2542	95298162 (5083, 5084)	Novel Protein sim. GBank gi 5225320 gb A040850.1 AF08310 sirutin type 2 [Homo sapiens]	Contains protein domain (PF00220) - Neurotrophylsial hormones, N- terminal Domain	UNCLASSIFIED	284488, 18108394, 52646385, 52646842, 65274572, 22278994, 56969386, 22278996, 284259, 52645080, 29331822, 29331824, 29331827, 35698902, 33656970, 264907, 284909, 52644045, 284510, 265008, 284512, 265007, 265008, 265009, 264910, 60431735, 52646317, 52644296, 265010, 265011, 265018, 265019, 18108351, 264683, 264285, 264685, 264687, 52644229, 264769, 21906786, 21906767, 21906765, 52644150, 33657023, 33657109, 52645129, 33657182, 27468261, 27468294, 33657349, 33695763, 18108374, 35698423, 35695855, 264631, 264634, 264635, 264595, 63373044, 18108385, 18108387, 67166518, 264583, 264584, 264585, 18108386, 264586, 65274572, 56183575, 22278999, 264259, 29331826, 264907, 264510, 264511, 264592, 284595, 264784, 264369, 264288, 264684, 264788, 264889, 21906765, 21906767, 21906769, 60170615, 264692, 264693, 55811576, 65274791, 264636, 264556, 18108381, 60170394, 264635, 18108385, 60432113, 22279000
2543	94139088 (5085, 5086)	Novel Protein sim. GBank gi 5419857 emb CAB46374.1 - (AL066723) hypothetical protein [Homo sapiens]	Contains protein domain (PF00076) - RNA recognition motif (a.k.a. RRM, RBD, or RNP domain)	UNCLASSIFIED	29331826, 264907, 264510, 264511, 264592, 284595, 264784, 264369, 264288, 264684, 264788, 264889, 21906765, 21906767, 21906769, 60170615, 264692, 264693, 55811576, 65274791, 264636, 264556, 18108381, 60170394, 264635, 18108385, 60432113, 22279000
2544	94218546 (5087, 5088)	Novel Protein sim. GBank gi 2498110 gb O63191 AEGP_RAT - APICAL ENDOSONAL GLYCOPROTEIN PRECURSOR	Contains protein domain (PF00629) - MAM domain.	UNCLASSIFIED	18108397, 52646355, 22278997, 264259, 60432049, 29331822, 29331825, 29331826, 29331827, 29331828, 264905, 264908, 265006, 265007, 265008, 87168559, 265017, 265018, 265019, 18108351, 264448, 264686, 264687, 264688, 21906765, 265020, 265021, 181083310, 18108374, 18108376, 18108381, 18108382, 18108383, 18108384, 18108385, 264482, 264583, 264587, 29331825, 264906, 265009, 60170831, 265017, 264369, 21906767, 60170615, 264692, 33657109
2545	8742845 (5089, 5090)	Novel Protein sim. GBank gi 33270460 gb A031591 - (AB014516) KIA0516 protein [Homo sapiens]	Contains protein domain (PF00097) - Zinc finger, C3HC4 type (RING finger)	UNCLASSIFIED	29331824, 265007, 22279002
2546	88093861 (5091, 5092)	Novel Protein sim. GBank gi 2969032 AF054586 - brain finger protein [Rattus norvegicus]	Contains protein domain (PF00097) - Zinc finger, C3HC4 type (RING finger)	UNCLASSIFIED	29331824, 265007, 22279002

2547	84143865 (5093, 5094)	Novel Protein sim. GBank g14929607/gbAJO40464_1 (AF151827) protein [Homo sapiens]	Contains protein domain (PF00183) - transport Mitochondrial carrier proteins	264488, 18108394, 52646842, 18106397, 56182575, 22778995, 56994075, 22278996, 22278997, 22778998, 264459, 26331822, 29331824, 29331826, 60432289, 29331827, 35690052, 29331828, 264404, 264508, 264905, 264906, 264908, 56712502, 264899, 56182493, 265006, 265007, 264512, 265008, 265009, 60170831, 60432228, 60451735, 265010, 265011, 265012, 264898, 265013, 265014, 264900, 264901, 265018, 265019, 264900, 264901, 265018, 265019, 264760, 18108351, 264682, 264448, 264288, 264369, 264684, 264685, 264687, 56181562, 264688, 264689, 21906765, 21906766, 21906767, 21906768, 29148627, 21906769, 55611957, 255020, 265021, 265022, 264690, 264691, 18108362, 264692, 264693, 27486281, 18108370, 18108374, 55810764, 55611578, 35686423, 35686455, 264635, 264636, 264655, 264637, 265381, 264537, 18108380, 264638, 56182323, 264598, 264599, 83373904, 18108385, 87188518, 22279002, 264564, 264566, 264567, 18108386, 22278999, 35686296, 56994075, 22278997, 22278999, 264259, 29331822, 29331824, 29331825, 29331826, 60432289, 29331827, 29331828, 56182435, 264611, 265007, 264512, 60433356, 87168559, 264684, 264369, 52644239, 265021, 33657023, 264632, 18108374, 52644332, 264557, 18108380, 18108381, 18108382, 18108384, 18108385, 60432113, 22279002, 22279002, 264583, 264587
2548	8879077 (5095, 5096)		UNCLASSIFIED	264488, 18108394, 52646842, 18106397, 56182575, 22778995, 56994075, 22278996, 22278997, 22778998, 264459, 26331822, 29331824, 29331826, 60432289, 29331827, 35690052, 29331828, 264404, 264508, 264905, 264906, 264908, 56712502, 264899, 56182493, 265006, 265007, 264512, 265008, 265009, 60170831, 60432228, 60451735, 265010, 265011, 265012, 264898, 265013, 265014, 264900, 264901, 265018, 265019, 264760, 18108351, 264682, 264448, 264288, 264369, 264684, 264685, 264687, 56181562, 264688, 264689, 21906765, 21906766, 21906767, 21906768, 29148627, 21906769, 55611957, 255020, 265021, 265022, 264690, 264691, 18108362, 264692, 264693, 27486281, 18108370, 18108374, 55810764, 55611578, 35686423, 35686455, 264635, 264636, 264655, 264637, 265381, 264537, 18108380, 264638, 56182323, 264598, 264599, 83373904, 18108385, 87188518, 22279002, 264564, 264566, 264567, 18108386, 22278999, 35686296, 56994075, 22278997, 22278999, 264259, 29331822, 29331824, 29331825, 29331826, 60432289, 29331827, 29331828, 56182435, 264611, 265007, 264512, 60433356, 87168559, 264684, 264369, 52644239, 265021, 33657023, 264632, 18108374, 52644332, 264557, 18108380, 18108381, 18108382, 18108384, 18108385, 60432113, 22279002, 22279002, 264583, 264587
2549	94106952 (5097, 5098)	Novel Protein sim. GBank g1728837spP39194ALU2_HUMAN - III ALU SUBFAMILY SQ WARNING ENTRY III	Contains protein domain (PF100412) - struct LIM domain containing proteins	56182575, 22278995, 22278997, 22278998, 22278999, 264259, 264500, 264406, 29331830, 265009, 265010, 265018, 264698, 21906764, 21906765, 21906766, 21906767, 21906769, 265020, 265021, 52644150, 264598, 264599, 83373904, 18108385, 35686423, 56182323, 18108382, 264587, 56182575, 26331827, 264405, 264512, 18108351, 56695917, 264637, 264638
2550	87778584 (5099, 5100)	Novel Protein sim. GBank g17143856 (p1432323 - nucleoporin p62 homolog - rat (fragment))	UNCLASSIFIED	264488, 18108394, 52646842, 18106397, 56182575, 22778995, 56994075, 22278996, 22278997, 22778998, 264459, 26331822, 29331824, 29331826, 60432289, 29331827, 35690052, 29331828, 264404, 264508, 264905, 264906, 264908, 56712502, 264899, 56182493, 265006, 265007, 264512, 265008, 265009, 60170831, 60432228, 60451735, 265010, 265011, 265012, 264898, 265013, 265014, 264900, 264901, 265018, 265019, 264760, 18108351, 264682, 264448, 264288, 264369, 264684, 264685, 264687, 56181562, 264688, 264689, 21906765, 21906766, 21906767, 21906768, 29148627, 21906769, 55611957, 255020, 265021, 265022, 264690, 264691, 18108362, 264692, 264693, 27486281, 18108370, 18108374, 55810764, 55611578, 35686423, 35686455, 264635, 264636, 264655, 264637, 265381, 264537, 18108380, 264638, 56182323, 264598, 264599, 83373904, 18108385, 87188518, 22279002, 264564, 264566, 264567, 18108386, 22278999, 35686296, 56994075, 22278997, 22278999, 264259, 29331822, 29331824, 29331825, 29331826, 60432289, 29331827, 29331828, 56182435, 264611, 265007, 264512, 60433356, 87168559, 264684, 264369, 52644239, 265021, 33657023, 264632, 18108374, 52644332, 264557, 18108380, 18108381, 18108382, 18108384, 18108385, 60432113, 22279002, 22279002, 264583, 264587

2551	95508400 (5101, 5102)	Novel Protein sim. G.Bank g/4337103g/AD18079) - (AF129756) NG26 [Homo sapiens]	Contains protein domain (PF00551) - UNCLASSIFIED	15103396, 55274572, 2278955, 2278998, 2278999, 264259, 29331822, 29331824, 29331825, 60432386, 29331826, 29331827, 264905, 56182435, 265007, 60433438, 55812038, 21006754, 65274444, 265017, 265018, 264605, 265019, 264288, 21006766, 21906768, 21006769, 265020, 60170615, 264993, 33657109, 39996423, 264938, 56182323, 53773044, 22779000
2552	95532620 (5103, 5104)		UNCLASSIFIED	56182375, 35696286, 29331824, 29331826, 29331827, 29331828, 29331829, 29331830, 29331831, 29331832, 56182435, 265007, 264991, 33109864, 264760, 55811957, 35695917, 33657023, 33657109, 18108374, 5581576, 35696423, 35695855, 56182323, 294558
2553	95508243 (5105, 5106)	Novel Protein sim. G.Bank g/17116585p/94797/T10_MOUSE - SER/THR-RICH PROTEIN T10 IN DGCR REGION	UNCLASSIFIED	264686, 264488, 263575, 264768, 29331826, 35696032, 35696423, 264601, 264511, 264602, 264910, 264634, 264760, 264555, 264762, 264906, 264592, 264691, 264566, 264908, 264684, 264567, 264909, 264766
2554	87761320 (5107, 5108)	Novel Protein sim. G.Bank g/1728835p/939162/ALUS_HUMAN - III ALU SUBFAMILY SC WARNING ENTRY III	cadherin	22278997, 29331822, 264508, 21900769, 33657023, 33657109, 56182323
2555	97627351 (5109, 5110)	Novel Protein sim. G.Bank g/4864319jemb/CA643260.T1 - ALU360841 hypothetical protein [Homo sapiens]	nuclease	29331824, 263972
2556	97645533 (5111, 5112)	Novel Protein sim. G.Bank g/106594 (NC003036) - RQ0923_1 [Homo sapiens]	Contains protein domain (PF00514) - UNCLASSIFIED Annulobacter-cadherin-like repeats	22278998, 264509, 33657102, 264683, 264684, 264685, 264686, 3367023, 33657109, 35695855, 264558, 264567
2557	99437803 (5113, 5114)	Novel Protein sim. G.Bank g/1191104p/P03211EBN1_EBV - EBNA-1 NUCLEAR PROTEIN	UNCLASSIFIED	264565
2558	97617591 (5115, 5116)	Novel Protein sim. G.Bank g/1505139jemb/CA639619.T1 - (AL049481) AIG1-like protein [Arabidopsis thaliana]	UNCLASSIFIED	22278997, 29331824, 6871417, 29331825, 264906, 264511, 265018, 264448
2559	88096382 (5117, 5118)	Novel Protein sim. G.Bank g/1505139jemb/CA639619.T1 - (AL049481) AIG1-like protein [Arabidopsis thaliana]	UNCLASSIFIED	22278997, 29331822, 29331828, 60433356, 265011, 264288, 264765, 264766, 264769
2560	87994530 (5119, 5120)	Novel Protein sim. G.Bank g/1505139jemb/CA644665.T1 - (AL078630) 573K4.3 (nm17M1-4) (novel 7 transmembrane receptor (rhodopsin family) (olfactory receptor LINE) protein) [Mus musculus]	Contains protein domain (PF00001) - 7 transmembrane receptor (rhodopsin family)	21906765, 21906766, 60432113, 264482
2561	8876675 (5121, 5122)	Novel Protein sim. G.Bank g/1505139jemb/CA644665.T1 - (AF044953) Nucleoside diphosphate kinase P-GV subunit [Homo sapiens]	UNCLASSIFIED	22278995, 35696286, 2278998, 2278997, 22278998, 2278999, 264259, 29331822, 29331824, 29331826, 29331827, 29331828, 29331829, 29331830, 29331831, 29331832, 56182435, 265007, 264991, 33109864, 264760, 55811957, 35695917, 265020, 60170615, 264993, 33657109, 18108374, 264634, 264558, 18103396, 53773044, 22779002

2580	88168788 (5157, 5160)	Novel Protein sim. GBank g123486528 (AC003080) - Similar to KIAA0299, 80% similarity to AB002297 (PID g224539) [Homo sapiens]			265007, 265018, 264762
2581	87595046 (5161, 5182)	Novel Protein sim. GBank g14066422p/AL202481 - (AF131809) [Unknown Homo sapiens]	Contains protein domain (PF00059) - collagen PDZ domain (Also known as DHR or GLGF).		56594075, 26331824, 26331826, 26331828, 264756, 87168559, 21900769, 265022, 35695855, 263981
2582	87786789 (5163, 5164)	Novel Protein sim. GBank g12733957 (AC005505) - putative phosphatidylinositol-4-phosphate 5-kinase [Arabidopsis thaliana]	eph		264488, 264907, 264908, 264910, 264764, 264684, 264766, 264636, 264555, 264565
2583	91220650 (5166, 5160)	Novel Protein sim. GBank g14378112p/CA116621.11 - (AL021579) J44G12.2 (similar to transcription factor RBP-L) [Homo sapiens]	Contains protein domain (PF00047) - transcription factor immunoglobulin domain		56161686, 264259, 264510, 264512, 264591, 264592, 264593, 264594, 264595, 264596, 264603, 264629, 5510764, 264630, 264637, 264565
2584	8043094 (5167, 5169)		UNCLASSIFIED		264910, 264768, 264693, 1818374, 55811576, 56182323
2585	80436126 (5168, 5170)	Novel Protein sim. GBank g12736151 (AF021639) - myotonic dystrophy kinase-related Cdc42-binding kinase [Rattus norvegicus]	kinase		264768
2586	91226136 (5171, 5172)				22278988, 264259, 26331822, 26331824, 26331827, 26331829, 264906, 265007, 265009, 264591, 60433356, 33657402, 265018, 264762, 264288, 21900766, 21906767, 21900769, 265022, 264891, 83373044, 55252486, 22279502
2587	86430942 (5173, 5174)				264908, 265019, 264768, 264693, 55811576, 56182323
2588	8074384 (5175, 5176)		UNCLASSIFIED		264584
2589	86515807 (5177, 5178)	Novel Protein sim. GBank g13021598p/CA171415 - (Y1088) nuclear protein [Xenopus laevis]	UNCLASSIFIED		35696032, 264906, 264908, 264907, 264908, 264909, 265018, 264767, 35698423, 265018
2590	87054526 (5179, 5180)	Novel Protein sim. GBank g12704689 (J02793) - alpha glucosidase II, alpha subunit [Mus musculus]	Contains protein domain (PF01055) - glucosylase Glycosyl hydrolases family 31		22278995, 26331830, 265006, 265010, 265017, 264639
2591	94392167 (5181, 5182)	Novel Protein sim. GBank g15702202p/AL047199.1/AF12916 - (AF129166) long-chain acyl-CoA synthetase 5 [Homo sapiens]	eph		264259, 26331822, 264108, 264908, 56182435, 265007, 265008, 33109984, 264448, 55811957, 265020, 1818370, 55811576, 22279002

[illegible]

2607	8762742 (5213, 5214)	Novel Protein sim. GBank g9426626(gjAAD30202.1) - (AF15022) mediator [Homo sapiens]		25331822, 26331825, 26331826, 26331827, 26331828, 264906, 264907, 264908, 66712502, 264926, 58182435, 55812038, 265010, 265017, 265018, 265019, 264768, 264689, 21906765, 55811987, 265020, 265022, 264692, 33657023, 264693, 33657109, 18108370, 264639, 56182323, 264631, 264907, 264909, 264692, 264759, 264639
2608	87134766 (5215, 5216)	Novel Protein sim. GBank g12226005 (U49973) - ORF2 function unknown [Homo sapiens]	Contains protein domain (PF00650) - histone	254488, 6527457, 3586248, 22278997, 22278999, 60432049, 264539, 56182161, 25331824, 26331825, 26331826, 26331827, 264107, 264108, 264905, 60433358, 60433359, 55812038, 265011, 60433358, 60433359, 55812038, 265011, 87168559, 265017, 265018, 264448, 264765, 264288, 264766, 264689, 21906765, 21906767, 21906769, 265020, 265021, 264691, 264692, 33657109, 27486281, 18108370, 65274791, 264636, 264556, 56182323, 18108385, 55526485
2609	94643791 (5217, 5218)	Novel Protein sim. GBank g1024689(gjP365241726) - HUMAN - HYPOTHETICAL PROTEIN KIAA0268 (KIA0116)	histone deacetylase family	254488, 6527457, 3586248, 22278997, 22278999, 60432049, 264539, 56182161, 25331824, 26331825, 26331826, 26331827, 264107, 264108, 264905, 56182435, 264112, 265006, 265007, 265008, 265009, 60433358, 60433359, 265011, 87168559, 265017, 264448, 264682, 264764, 264288, 264766, 264689, 21906765, 21906767, 21906769, 265020, 265021, 264691, 264692, 33657109, 27486281, 18108370, 65274791, 264636, 264556, 56182323, 18108385, 55526485
2610	88177654 (5219, 5220)	Novel Protein sim. GBank g10336825(gjAAD77889) - (AF106473) leucine-rich-domain inter-acting protein 1; LeR inter-acting protein 1; LEAP1 [Mus musculus]	transcription factor	18108394, 22278994, 56994075, 60432049, 264259, 29331822, 26331825, 60432289, 29331827, 264107, 264108, 264905, 56182435, 264112, 265006, 265007, 265008, 265009, 60433358, 60433359, 265011, 87168559, 265017, 264448, 264682, 264764, 264288, 264766, 264689, 21906765, 21906767, 21906769, 265020, 265021, 264691, 264692, 33657109, 27486281, 18108370, 65274791, 264636, 264556, 56182323, 18108385, 55526485
2611	87428990 (5221, 5222)	Novel Protein sim. GBank g103876761(embCAA92994) - (268760) predicted using GeneIndex. Similarity to Mouse FKBP-type papilloma-prolyl isomerase	Contains protein domain (PF00254) - isomerase	22278996, 265017, 264684, 21906768, 22279000
2612	87771196 (5223, 5224)	Novel Protein sim. GBank g105679136(gjAAD6874) - (AF160934) FK506-binding protein [SWFKB3_MOUSE] [Caenorhabditis elegans]	transport	265009, 264910, 264759, 265017, 21906767, 18108365, 18108384, 60432113
2613	79481496 (5225, 5226)	Novel Protein sim. GBank g10533003(gjAAD5009) - (AF16118) P55T protein [Mus musculus]	UNCLASSIFIED	264685
2614	87643946 (5227, 5228)	Novel Protein sim. GBank g10533003(gjAAD5009) - (AF16118) P55T protein [Mus musculus]	Contains protein domain (PF00625) - Guanylate kinase	22278998, 22278999, 29331825, 264508, 264906, 21906754, 264602, 264766, 264769, 52644229, 21906765, 33657109, 27486264, 18108370, 263972, 264555, 60432113
2615	87181096 (5229, 5230)	Novel Protein sim. GBank g10533003(gjAAD5009) - (AF16118) P55T protein [Mus musculus]	UNCLASSIFIED	264768, 18108394, 264692, 264689, 21906765, 21906767, 21906769, 265020, 265021, 18108370, 264511, 264512, 264910, 264635, 264595, 265010, 264404, 264563, 264764, 264685, 264766

2616	87428895 (5231, 5232)	Novel Protein sim. GBank gi 3876761 emb CA329941 - (265760) predicted using GeneFinder. Similarity to Mouse FKBP-type peptidyl-prolyl isomerases	Contains protein domain (PF00254) - isomerases	22278965, 22278997, 22278998, 60432049, 60432289, 264268, 60433356, 264594, 60434388, 33109954, 87168474, 265011, 265017, 265019, 264288, 264766, 21906765, 21906767, 265020, 265021, 18108376, 18108377, 18108387, 87168518, 264482, 264587
2617	86578888 (5231, 5232)	Novel Protein sim. GBank gi 2585116 PF34884ALL1 - J WARNING ENTRY III	kinase	265010, 265019, 264369, 264693, 56611576, 22279002
2618	91231662 (5235, 5236)	Novel Protein sim. GBank gi 3319282 AF491031 - Huntingtin Interacting protein Homo sapiens	Contains protein domain (PF00397) - UNCLASSIFIED WW domain	264489, 22278998, 264591, 264259, 29331822, 264102, 264608, 264808, 264907, 66712502, 29331830, 265008, 264910, 265009, 60433356, 60433438, 264756, 21906754, 265011, 87168559, 265017, 265018, 264389, 264288, 264766, 264768, 264689, 21906765, 21906766, 21906767, 35696917, 265020, 265022, 33657023, 264692, 33657109, 264628, 18108374, 36696645, 18108381, 83373044, 18108385, 18108388, 56528886, 264463
2619	87640000 (5237, 5238)	Novel Protein sim. GBank gi 3431772 U664111 - putative type III alcohol dehydrogenase [Drosophila melanogaster]	Contains protein domain (PF00465) - dehydrogenase iron containing alcohol	284239, 60432288, 60433438, 21906754, 264369, 60432113, 324566
2620	95314841 (5239, 5240)	Novel Protein sim. GBank gi 3325670 p CAJ169211 - (AF090436) dactylusd variant 1 [Mus musculus]	Contains protein domain (PF00638) - UNCLASSIFIED PHO-finger	53841507, 53841515, 53845843, 65737672, 22278995, 56994075, 35696246, 22278998, 60432049, 264269, 59245900, 29331822, 29331824, 29331823, 29331826, 29331828, 35696052, 264907, 66712502, 265008, 60433356, 33657402, 52646317, 21906754, 87168474, 265010, 265017, 265018, 264448, 264389, 264288, 264687, 264768, 52644229, 264688, 264689, 21906765, 21906768, 35696917, 52644150, 264692, 33657109, 35696763, 35696423, 264556, 52644332, 18108382, 83373044, 18108385, 18108387, 65274727, 87168518, 60432113, 22279002, 264594, 264636
2621	86253408 (5241, 5242)	Novel Protein sim. GBank gi 2585116 PF34884ALL1 - J WARNING ENTRY III		264488, 264908, 264907, 264908, 264932, 265007, 264738, 35695917, 264634, 264636, 264583, 264462
2622	91730390 (5243, 5244)	Novel Protein sim. GBank gi 2585116 PF34884ALL1 - J WARNING ENTRY III		
2623	91635006 (5245, 5246)	Novel Protein sim. GBank gi 3880355 emb CA8052081 - (282285) predicted using GeneFinder [Caenorhabditis elegans]	UNCLASSIFIED	29331834, 35696052, 265007, 265010, 264288, 20148229

2624	61039300 (5247, 5248)	Novel Protein sim. GBank gll1880355[met](CA03239) - (232285) predicted using GeneFinder [Caenorhabditis elegans]	UNCLASSIFIED	56151696, 22278996, 22278997, 22278998, 22278999, 264259, 50331622, 26331824, 56192181, 26331823, 60432289, 29331828, 35696052, 29146489, 66712502, 52644045, 265007, 265008, 60433356, 33109954, 21906754, 265010, 265011, 265019, 264448, 264288, 21906765, 21906766, 21906767, 29146629, 35695917, 265021, 265022, 27486265, 18108370, 80431528, 55811576, 35695955, 56182323, 18108385, 87166518, 22279002, 18108391
2625	86432068 (5249, 5250)	Novel Protein sim. GBank gll28874236[ob](BA424857) - (232285) predicted using GeneFinder [Homo sapiens]	UNCLASSIFIED	22278996, 265007, 265009, 264448, 21906767, 265021, 264558, 18108385, 22278997, 22278998, 264559, 604332048
2626	16533767 (5251, 5252)	Novel Protein sim. GBank gll6787419 (L20392) - scin flarant protein [Galus nigrus]	UNCLASSIFIED	29331822, 29331823, 29331825, 29331827, 35696052, 29331828, 284807, 264909, 265008, 264591, 60433355, 60433438, 265010, 265017, 265018, 264369, 264288, 18108357, 21906765, 21906768, 265022, 65274791, 264638, 18108387, 87166518, 22279002
2627	87636823 (5253, 5254)	Novel Protein sim. GBank gll8482[pr](A2707) - prolase-ich phosphoprotein [gerc PRH1, Dh. allele] - human	UNCLASSIFIED	29331822, 29331823, 29331825, 29331827, 35696052, 29331828, 284807, 264909, 265008, 264591, 60433355, 60433438, 265010, 265017, 265018, 264369, 264288, 18108357, 21906765, 21906768, 265022, 65274791, 264638, 18108387, 87166518, 22279002
2628	94848254 (5255, 5256)	Novel Protein sim. GBank gll1235352[met](CA118608) - (AL022578) dJ393P12.2 (hypothetical Proteine-ich protein KIAA0269 LIKE) [Homo sapiens]	UNCLASSIFIED	29331822, 29331823, 29331825, 29331827, 35696052, 29331828, 284807, 264909, 265008, 264591, 60433355, 60433438, 265010, 265017, 265018, 264369, 264288, 18108357, 21906765, 21906768, 265022, 65274791, 264638, 18108387, 87166518, 22279002
2629	87767490 (5257, 5258)	Novel Protein sim. GBank gll4929595[gh]AD04056 (JAF15182 - (AF151821) CGI-63 protein [Homo sapiens])	UNCLASSIFIED	29331822, 29331823, 29331825, 29331827, 35696052, 29331828, 284807, 264909, 265008, 264591, 60433355, 60433438, 265010, 265017, 265018, 264369, 264288, 18108357, 21906765, 21906768, 265022, 65274791, 264638, 18108387, 87166518, 22279002
2630	79184364 (5259, 5260)	Novel Protein sim. GBank gll1235352[met](CA118608) - (AL022578) dJ393P12.2 (hypothetical Proteine-ich protein KIAA0269 LIKE) [Homo sapiens]	UNCLASSIFIED	29331822, 29331823, 29331825, 29331827, 35696052, 29331828, 284807, 264909, 265008, 264591, 60433355, 60433438, 265010, 265017, 265018, 264369, 264288, 18108357, 21906765, 21906768, 265022, 65274791, 264638, 18108387, 87166518, 22279002
2631	84455909 (5261, 5262)	Novel Protein sim. GBank gll1235352[met](CA118608) - (AL022578) dJ393P12.2 (hypothetical Proteine-ich protein KIAA0269 LIKE) [Homo sapiens]	UNCLASSIFIED	29331822, 29331823, 29331825, 29331827, 35696052, 29331828, 284807, 264909, 265008, 264591, 60433355, 60433438, 265010, 265017, 265018, 264369, 264288, 18108357, 21906765, 21906768, 265022, 65274791, 264638, 18108387, 87166518, 22279002
2632	38720414 (5263, 5264)	Novel Protein sim. GBank gll1235352[met](CA118608) - (AL022578) dJ393P12.2 (hypothetical Proteine-ich protein KIAA0269 LIKE) [Homo sapiens]	UNCLASSIFIED	29331822, 29331823, 29331825, 29331827, 35696052, 29331828, 284807, 264909, 265008, 264591, 60433355, 60433438, 265010, 265017, 265018, 264369, 264288, 18108357, 21906765, 21906768, 265022, 65274791, 264638, 18108387, 87166518, 22279002

2638	94326733 (5275, 5276)	Novel Protein sim. GBank g1429669[gb]AD34.105.1 [AF15186 - (AF15186) CGI-110 protein (Homo sapiens)]	Contains protein domain (PF00076) - RNA recognition motif. (p.k.a. RRM, RBD, or RNP domain)	UNCLASSIFIED	60424178, 5264507, 52646842, 18103398, 56182575, 22278995, 22278996, 35698286, 29331822, 60424263, 29331826, 35696052, 29146488, 264805, 52644045, 56182435, 60433396, 33857402, 55812038, 55811386, 265019, 264288, 264709, 52644223, 56181562, 29148827, 29148829, 55811857, 33857023, 62274820, 33857109, 35695763, 18108374, 55810764, 35696423, 55811376, 35695955, 60431690, 56182523, 60432713, 264468
2639	95361346 (5277, 5278)	Novel Protein sim. GBank g12190007[gb]A3303561 - (A3004100) phosphotyrosine synthase II [Coccidius glaucus]		synthase	264468, 29331825, 35696052, 264508, 264509, 264805, 264512, 33857402, 60433438, 264708, 85658542, 264600, 265020, 265021, 33857109, 264628, 35698423, 264565, 264639, 264563, 264554, 264565, 264566, 264486
2640	87781330 (5279, 5280)	Novel Protein sim. GBank g1315818 (AF067617) - contains similarity to chromo (chromatin organization modifier) domains (Pfam: chromo.tmm. score: 17.76 and 27.94) and to helicases conserved C-terminal domain (Pfam: helicase_C.tmm. score: 87.00) [Caenorhabditis elegans]	Contains protein domain (PF00271) - Helicases conserved C-terminal domain	helicase	29331822, 29331826, 264606, 33103954, 265017, 265019, 21905768, 35698783, 264636, 264637, 18108387
2641	11659834 (5281, 5282)	Novel Protein sim. GBank g12564955 (AF030001) - unknown [Mus musculus]		UNCLASSIFIED	264828, 29331822, 33857402, 265019, 265020, 265021, 33857109, 264628, 264629, 264630, 264631, 264634, 35696486, 22278997
2643	87643861 (5285, 5286)	Novel Protein sim. GBank g1440304[emb]C4538793.11 - (AL055878) putative protein [Arabidopsis thaliana]	Contains protein domain (PF00270) - DEAD/DEAH box helicase	-helicase	22278997, 264259, 29145459, 56182435, 264910, 265010, 18108351, 264682, 264683, 264369, 264684, 264685, 264686, 29148827, 264690, 33857109, 18108370, 265073, 264682, 18108374, 264634, 264557, 264558, 18108385, 264482
2644	88177671 (5287, 5288)	Novel Protein sim. GBank g1378979[gb]AA07502.11 - (AF065869) actin binding protein MAYVEN [Homo sapiens]	Contains protein domain (PF06651) - BTB/POZ domain	nuc1_rept	264107, 264687
2645	17277228 (5289, 5290)	Novel Protein sim. GBank g11706722[sp]P49749[ev]2. MOUSE - HOMEODOMAIN EVEN SKIPPED HOMOLOG PROTEIN 2 (EVA-2)		UNCLASSIFIED	265007
2648	94148542 (5291, 5292)	Novel Protein sim. GBank g11706722[sp]P49749[ev]2. MOUSE - HOMEODOMAIN EVEN SKIPPED HOMOLOG PROTEIN 2 (EVA-2)		UNCLASSIFIED	264899, 264687, 264632, 83373044

2647	91212378 (5294, 5294)			UNCLASSIFIED	56182515, 22276966, 35662686, 22276966, 18182122, 56182101, 26331823, 60424269, 60432749, 56182101, 26331823, 2645008, 265007, 55612038, 33109594, 21906764, 33657004, 265019, 264448, 264288, 56181562, 21906765, 21906766, 21906768, 21906769, 35655917, 265020, 265021, 52644150, 264693, 33657109, 33657349, 60431526, 18108374, 55810764, 35696423, 56182323, 60432113, 22276902, 264564
2648	87600387 (5295, 5296)				2816498, 56182434, 33109654, 265011, 264692, 55811957, 35695917, 264690, 263976, 18103377, 35696423, 60432113, 56182375, 35096286, 22276998, 26331824, 26331825, 60432289, 60712302, 56182435, 21906765, 21906766, 21906768, 21906769, 35655917, 265020, 21906765, 21906766, 60712302, 56182435, 263987, 18108370, 263976, 6010394, 60432113, 22276902, 264693, 264695
2649	94128783 (5297, 5298)	Novel Protein sim. GBank g10041882 (A004339) - unknown function, similar to Y09105 (P1D-3166611) [Homo sapiens]		UNCLASSIFIED	23331824, 26331826, 26331827, 265007, 55812038, 21906754, 18108366, 18108384, 22276902, 264567
2650	97297333 (5300, 5300)	Novel Protein sim. GBank g10360271 (p01BAA431908.1) - (AB029353) HRF-1 [3] (Drosophila melanogaster)	Contains protein domain (PF00054) - synthase Laminin G domain		264692
2651	98088745 (5301, 5302)	Novel Protein sim. GBank g1042022 (p01BAA47409.1) - (AB020879) KIA0085 protein [Homo sapiens]		UNCLASSIFIED	264692
2652	10433725 (5303, 5304)	Novel Protein sim. GBank g10449386 (p01BAA481123.1) - (Z98551) predicted using hcdEon: MAL396.26 (PFC0845c), Hypothetical protein, len: 167 aa. Similarity to model organism hypothetical proteins (C. elegans, D. melanogaster, S. cerevisiae & S. pombe) C. elegans protein ZC987.5 (TR...		UNCLASSIFIED	265018, 18108370, 18108387, 264566
2653	97736735 (5305, 5306)				
2654	95103240 (5307, 5308)				60424178, 65274372, 56182515, 264659, 56182181, 264604, 56182435, 55811957, 35695917, 265021, 263976, 55810764, 95274791, 56182233, 83373044, 65274727, 56182575, 56181686, 264082, 264258, 56182181, 50432289, 264607, 33657402, 55812038, 21906754, 87166559, 265017, 264448, 264369, 264488, 21906765, 21906766, 21906767, 21906768, 33657109, 18108370, 264628, 55811576, 264558, 264639, 83373044, 56526486, 264404, 96427113, 264693
2655	91226018 (5309, 5310)	Novel Protein sim. GBank g10367327 (p01BAA30286.1) - (Z11051) predicted using GeneRuler, similar to Zinc finger, C2HC4 type (RING finger), cDNA EST Y443K5.3 comes from Homo sapiens, cDNA EST Y443K5.3 comes from the gene [Caenorhabditis elegans]	Contains protein domain (PF00087) - Zinc finger, C2HC4 type (RING finger)	transcription factor	
2656	94562601 (5311, 5312)	Novel Protein sim. GBank g1004371 (p01BAA35523) - (AB011189) KIA0597 protein [Homo sapiens]			

[illegible]

2668	91215716 (5335, 5336)	Novel Protein sim. GBank gi 4544146 ref nr_068327.1 p270; ZYG homolog	UNCLASSIFIED	56181688, 35869288, 2278988, 2278988, 2278988, 56182781, 29331024, 60424289, 29331825, 35890502, 29331828, 60712502, 56182435, 60433396, 264758, 21906754, 55811386, 265011, 87168559, 265017, 265019, 55811150, 264448, 264389, 264288, 21906765, 21905766, 21906768, 55811857, 35895917, 265020, 265021, 33657023, 264692, 33657109, 36655763, 60431528, 18108374, 35096423, 55811576, 2646534, 60431650, 33373044, 18108385, 87168519, 22270000, 264593, 264564
2669	85415721 (5337, 5338)	Novel Protein sim. GBank gi 21470172 ref JC4589 - proline rich protein - rat		254489, 264689, 21905767, 63274572, 56182575, 21906768, 2318627, 21906769, 23146029, 35895906, 265017, 265019, 55811857, 33657023, 264692, 33657109, 36655763, 60431528, 18108374, 35096423, 55811576, 2646534, 60431650, 33373044, 18108385, 87168519, 22270000, 264593, 264564
2670	87613234 (5339, 5340)	Novel Protein sim. GBank gi 1735323 ref Q1082YD8B_SCHPO - HYPOTHETICAL 94.9 KD PROTEIN C2E12.1TC IN CHROMOSOME I	Contains protein domain (PF00623) - PHD-finger	254489, 264689, 21905767, 63274572, 56182575, 21906768, 2318627, 21906769, 23146029, 35895906, 265017, 265019, 55811857, 33657023, 264692, 33657109, 36655763, 60431528, 18108374, 35096423, 55811576, 2646534, 60431650, 33373044, 18108385, 87168519, 22270000, 264593, 264564
2671	91214958 (5341, 5342)	Novel Protein sim. GBank gi 4788277 ref AD029444 - (AF06425) very long-chain acyl-CoA synthetase homolog 2, VLCS-12 [Homo sapiens]	transport	55811857, 33657023, 264692, 33657109, 36655763, 60431528, 18108374, 35096423, 55811576, 2646534, 60431650, 33373044, 18108385, 87168519, 22270000, 264593, 264564
2672	87598123 (5343, 5344)	Novel Protein sim. GBank gi 1965348 ref AD04677 - (AC006341) - Contains two PF01344 Katch motif domains [Arabidopsis thaliana]	Contains protein domain (PF01344) - Katch motif	55811857, 33657023, 264692, 33657109, 36655763, 60431528, 18108374, 35096423, 55811576, 2646534, 60431650, 33373044, 18108385, 87168519, 22270000, 264593, 264564

2673	87430749 (5345, 5346)	Novel Protein sim. G3ank gll5457337[ambjCAB41505.2] - (AJ236876) poly(ADP-ribosyl) polymerase-2 (Homo sapiens)	Contains protein domain (PF00544) - polymerase catalytic region.	22278995, 22278996, 22278997, 22278998, 29331822, 29331824, 29331826, 35696052, 60433438, 87168474, 87168559, 265017, 265018, 265019, 264448, 21906768, 21906769, 265020, 265021, 33657109, 27488282, 35995765, 60431850, 60170394, 87168518, 264563
2674	94847721 (5347, 5348)	Novel Protein sim. G3ank gll473624iprNF_004280.1[hnRF3 - nuclear factor (erythroid-derived 2)-like 3]	Contains protein domain (PF00770) - transcription factor	264488, 22278996, 35696286, 264091, 264499, 29331824, 29331826, 35696052, 264311, 50812038, 50835442, 264786, 21906765, 35995917, 264629, 35696423, 81083383, 87108516
2675	79563335 (5349, 5350)		UNCLASSIFIED	264566, 265008
2676	79528393 (5351, 5352)		UNCLASSIFIED	18108394, 65274372, 56182575, 22278994, 22278995, 56994072, 22278996, 35696286, 22278997, 22278998, 264490, 264259, 52645080, 29331824, 29331825, 60432289, 29331827, 29331828, 35696052, 29145499, 29331830, 264908, 52644545, 265006, 265007, 265008, 265009, 60432229, 60433356, 60433438, 53812038, 265010, 265011, 87168559, 265017, 265019, 18108351, 264682, 264448, 264483, 264288, 21906765, 21906766, 21906767, 21906768, 21906769, 53811957, 205020, 265021, 60170613, 52644150, 264691, 33657023, 33657024, 35696052, 21906768, 21906769, 33657025, 264908, 52644545, 265006, 265007, 55811576, 35696423, 35696455, 18108377, 55811576, 35696423, 35696455, 83373044, 18108387, 22279000, 22279002, 264564
2677	94329600 (5353, 5354)	Novel Protein sim. G3ank gll10730042ipr[S2154 - acetyl-CoA synthetase - fruit fly (Drosophila melanogaster)]	Contains protein domain (PF10001) - synthase AMP-binding enzyme	

2704	87549515 (5407, 5408)	Novel Protein sim. GBank g[133569] (p[NA1863294] - (AF008556) Implantation-associated protein [Ratida nonvegca])		254488, 22278985, 22278986, 29331828, 254489, 264500, 264501, 264502, 264503, 264504, 264505, 264511, 2645702, 264620, 264602, 265017, 264605, 264761, 18108351, 264764, 264687, 264769, 265021, 264691, 264692, 18108362, 264693, 18108370, 18108374, 264634, 264635
2705	87771745 (5408, 5410)			254489, 264509, 264511, 264512, 264510, 264593, 87768474, 264604, 264688, 264687, 264769, 264638, 264656, 264486
2706	94326786 (5411, 5412)	Novel Protein sim. GBank g[326592] (emb[CAAT16821.1] - (AL021728) /prediction(method: match) (desc: [Drosophila melanogaster])	UNCLASSIFIED	254488, 52646842, 65274572, 22278994, 56964075, 22278997, 254459, 29331824, 29331825, 29331826, 29331828, 33656970, 264907, 264908, 264909, 52644045, 56182435, 265000, 265007, 60433438, 58672038, 21900754, 53644296, 265010, 254601, 265017, 265019, 264681, 264446, 264682, 264683, 264684, 264685, 21908766, 21906769, 55811957, 36595917, 265020, 265021, 60170815, 264690, 264691, 33657033, 264692, 264693, 68274620, 27486284, 263972, 18108374, 18108377, 264635, 264636, 264556, 60170394, 83373044, 65274727, 87186518, 22279000, 22278996, 22278998, 56182435, 21900754, 87186559, 265017, 264448, 52645129
2707	86098939 (5413, 5416)	Novel Protein sim. GBank g[3417284 (AC004381) - (Unknown gene product [Homo sapiens])		22278996, 22278998, 56182435, 21900754, 87186559, 265017, 264448, 52645129
2708	91011351 (5415, 5416)	Novel Protein sim. GBank g[545790] (b[14778] - DAAPP, 32-diopamine and cAMP-regulated phosphoprotein [human, brain, Peptide, 204 aa])	UNCLASSIFIED	65274572, 264629, 29331822, 29331823, 6042289, 29331826, 29331827, 29331828, 294909, 264510, 265007, 264910, 60433596, 6043498, 33108954, 265010, 265011, 254389, 264288, 264765, 264693, 264695
2709	94453988 (5417, 5418)	Novel Protein sim. GBank g[3165705 (AC004780) - F17127.1 [Homo sapiens])		29331822, 18108370, 18108374, 83373044
2710	87627973 (5419, 5420)	Novel Protein sim. GBank g[446531] (m[JC4337928] - (AL031432) d4465N02.2.1 (PUTATIVE novel protein) (isoform 1) [Homo sapiens])	UNCLASSIFIED	29331824, 264759, 264693, 18108382, 18108388

2711	94111920 (5421, 5422)	Novel Protein sim. GBank gl3122400sp035682/MUC_MOUSE - MYELOID UPREGULATED PROTEIN	284488, 264687, 526455, 264789, 21906764, 21906765, 21906767, 21906768, 21906769, 55811957, 56994075, 22278997, 22278998, 265020, 265021, 264690, 264259, 264691, 264692, 33657023, 29331822, 29331824, 00424289, 29331826, 33657182, 29331827, 27468282, 33657349, 264508, 264905, 264907, 00431528, 264909, 264909, 55810764, 35696423, 05274791, 35659555, 265007, 264910, 00431630, 00432225, 55810765, 55810766, 55810767, 55810768, 18108374, 51060754, 33657024, 17168418, 17168419, 17168420, 60432113, 265017, 22275000, 265018, 265019, 22275002, 22275000, 265018, 265019, 22275002, 264760, 55811150, 264691, 18108351, 264555, 264764, 264566, 264288, 264766
2712	94372071 (5423, 5424)	Novel Protein sim. GBank gl5081315/gblAAD39343, jf407660 - (AF076607) predilectic NOD sera-reactive autoantigen [Mus musculus]	Contains protein domain (PF00515) - TPR Domain 284488, 35696286, 22275998, 264259, 29331824, 60432283, 35696092, 264508, 264906, 66712502, 526455, 265006, 60432229, 33657402, 60433356, 265010, 285019, 18108351, 264681, 264288, 264685, 21906765, 21906766, 21906768, 21906769, 55811957, 35695917, 265020, 265021, 6070815, 264691, 264692, 33657023, 33657109, 33657182, 33657349, 18108370, 18108374, 35696423, 35696555, 264555, 264760, 55811150, 264691, 18108351, 66714117, 264906, 264691
2713	86003064 (5425, 5426)	Novel Protein sim. GBank, gll2477513 (AC002398) - F25965_3 [Homo sapiens]	UNCLASSIFIED
2714	13528218 (5427, 5428)	Novel Protein sim. GBank, gll4321965/gblAAD158971 - (AF067430) Smarce1-related protein [Mus musculus]	UNCLASSIFIED
2715	94122454 (5429, 5430)	Novel Protein sim. GBank, gll4321965/gblAAD158971 - (AF067430) Smarce1-related protein [Mus musculus]	UNCLASSIFIED
2716	86003068 (5431, 5432)	Novel Protein sim. GBank, gll2477513 (AC002398) - F25965_3 [Homo sapiens]	UNCLASSIFIED
2717	80077461 (5433, 5434)	Novel Protein sim. GBank, gll3327046/gblAA315911 - (AB014156) KIAA0616 protein [Homo sapiens]	glycoprotein
2718	79604082 (5435, 5436)	Novel Protein sim. GBank, gll746495 (U23515) - weakly similar to gastrula zinc finger protein (Caenorhabditis elegans)	264693
2719	88180423 (5437, 5438)	Novel Protein sim. GBank, gll746495 (U23515) - weakly similar to gastrula zinc finger protein (Caenorhabditis elegans)	29331822, 87166459, 265019, 265021, 52644150, 264691

[illegible]

[illegible]

2735	87712338 (5468, 5470)	Novel Protein sim. GBank gi 350559 AC003278 - ESTs gbl121276, gbl145403, and gbl14586113 come from this gene. [Arabidopsis thaliana]	glycoprotein	22278996, 60437280, 20331827, 26146488, 2841008, 264900, 264112, 33657402, 87168474, 265017, 264762, 264448, 264784, 264684, 21906765, 264693, 33657105, 263976, 264638, 264630, 264657, 22278000, 22278002, 264587
2736	80247655 (5471, 5472)		UNCLASSIFIED	264905, 264628, 264629, 263978, 264632, 264564
2737	87804526 (5473, 5474)	Novel Protein sim. GBank gi 7555651 gb 54424898 (CS3850) hepatoma-derived from [Homo sapiens]		264488, 265009, 264788, 264691
2738	85731803 (5475, 5476)	Novel Protein sim. GBank gi 5420337 emb CA348679.1 - (A2434559) prokaryotic phosphoglycerate kinase [Leishmania major]	UNCLASSIFIED	264684, 63373044, 264566
2739	94318834 (5477, 5478)	Novel Protein sim. GBank gi 3417366 emb CAA754961 - (Y15197) microtubule-associated protein, MAP-115 [Mus musculus]	UNCLASSIFIED	264488, 56182575, 22278996, 35695386, 22278997, 22278998, 22278999, 2631822, 29331824, 29331825, 29331827, 35696092, 29331828, 29146468, 29331830, 265006, 265007, 265009, 60432229, 33657402, 55812038, 87168474, 265010, 265011, 265017, 285018, 265019, 284805, 264681, 264288, 294369, 52644229, 21906765, 21906766, 21906787, 21906788, 21906789, 265020, 265022, 264691, 264692, 33657109, 18108370, 18108374, 55810784, 35695585, 264634, 60431850, 264639, 56182323, 18108382, 18108385, 65274727, 22278002, 264564
2741	88047515 (5481, 5482)	Novel Protein sim. GBank gi 3242784 AC005154 - similar to protein U28928 (PID:p81308) [Homo sapiens]	UNCLASSIFIED	22278996, 52644046, 53644229, 21906768, 22278997, 22278998, 22278999, 2631822, 29331824, 29331825, 29331827, 35696092, 29331828, 29146468, 29331830, 265006, 265007, 265009, 60432229, 33657402, 55812038, 87168474, 265010, 265011, 265017, 285018, 265019, 284805, 264681, 264288, 294369, 52644229, 21906765, 21906766, 21906787, 21906788, 21906789, 265020, 265022, 264691, 264692, 33657109, 18108370, 18108374, 55810784, 35695585, 264634, 60431850, 264639, 56182323, 18108382, 18108385, 65274727, 22278002, 264564
2742	87648644 (5483, 5484)	Novel Protein sim. GBank gi 4758412 ref NP_004472.1 gb GALN - UDP-N-acetyl-alpha-D-galactosamine polypeptide N-acetylglucosaminyltransferase 2 [Gallus gallus]	UNCLASSIFIED	22278996, 52644046, 53644229, 21906768, 22278997, 22278998, 22278999, 2631822, 29331824, 29331825, 29331827, 35696092, 29331828, 29146468, 29331830, 265006, 265007, 265009, 60432229, 33657402, 55812038, 87168474, 265010, 265011, 265017, 285018, 265019, 284805, 264681, 264288, 294369, 52644229, 21906765, 21906766, 21906787, 21906788, 21906789, 265020, 265022, 264691, 264692, 33657109, 18108370, 18108374, 55810784, 35695585, 264634, 60431850, 264639, 56182323, 18108382, 18108385, 65274727, 22278002, 264564
2743	87627997 (5485, 5486)	Novel Protein sim. GBank gi 468337 emb CA837992 - (AL031432) DJ46SNG2.2.1 (PUTATIVE novel protein) (isoform 1) [Homo sapiens]	UNCLASSIFIED	22278996, 52644046, 53644229, 21906768, 22278997, 22278998, 22278999, 2631822, 29331824, 29331825, 29331827, 35696092, 29331828, 29146468, 29331830, 265006, 265007, 265009, 60432229, 33657402, 55812038, 87168474, 265010, 265011, 265017, 285018, 265019, 284805, 264681, 264288, 294369, 52644229, 21906765, 21906766, 21906787, 21906788, 21906789, 265020, 265022, 264691, 264692, 33657109, 18108370, 18108374, 55810784, 35695585, 264634, 60431850, 264639, 56182323, 18108382, 18108385, 65274727, 22278002, 264564
2744	94126300 (5487, 5488)	Novel Protein sim. GBank gi 3860433 emb CAA91399 - (226821) similar to mitochondrial RNA splicing MSR4 like protein, cDNA EST ENBL-009271 comes from this gene [Chromolaoba elegans]	UNCLASSIFIED	22278996, 52644046, 53644229, 21906768, 22278997, 22278998, 22278999, 2631822, 29331824, 29331825, 29331827, 35696092, 29331828, 29146468, 29331830, 265006, 265007, 265009, 60432229, 33657402, 55812038, 87168474, 265010, 265011, 265017, 285018, 265019, 284805, 264681, 264288, 294369, 52644229, 21906765, 21906766, 21906787, 21906788, 21906789, 265020, 265022, 264691, 264692, 33657109, 18108370, 18108374, 55810784, 35695585, 264634, 60431850, 264639, 56182323, 18108382, 18108385, 65274727, 22278002, 264564

2745	87740125 (6488, 5400)	Novel Protein sim. GBank (g4405798)(JAD1/8826) - (AF038863) RNA helicase [Homo sapiens]	Contains protein domain (PF00271) - Helicases conserved C-terminal domain	35656296, 264509, 264905, 264758, 264510, 264512, 264510, 264758, 264801, 265017, 264904, 264763, 264288, 264636, 264769, 264693, 35696423, 35695855, 264634, 264636, 264953, 264564, 264865
2746	95418601 (5491, 5492)	Novel Protein sim. GBank (g4758738)(eIFNP_004850, ItpMTA1 - metastasis associated 1	Contains protein domain (PF00320) - GATA zinc finger	22776996, 22776998, 22776999, 29331822, 29331826, 29331827, 35696052, 29331828, 264905, 264906, 264907, 264908, 264909, 52644045, 265006, 60170831, 264598, 55812036, 265018, 264683, 264686, 21906785, 21906787, 21906788, 21906789, 265019, 264905, 264906, 264907, 264908, 35657100, 18100386, 18100387, 264558, 18100385, 22776900, 264953
2747	94112677 (5493, 5494)	Novel Protein sim. GBank (g4557803)(eIFNP_000282, ItpNPC1 - Niemann Pick disease, type C1	glycoprotein	35656296, 52644007, 18100384, 22776999, 26504500, 29331824, 56182181, 29331826, 29331827, 35696052, 284907, 264908, 284909, 265009, 33109954, 55811396, 87168474, 265010, 87168559, 264603, 265019, 264760, 264686, 264768, 21906789, 35695917, 60170815, 264692, 33657023, 52645129, 27486284, 60431528, 18100374, 35696423, 35695855, 264595, 55192233, 18100385, 264482
2748	91214983 (5495, 5496)	Novel Protein sim. GBank (g4191272)(embCAG09884) - (J012295) ap4g protein (Rizidium aff)	Contains protein domain (PF00648) - P-loop domain	53274512, 29331828, 264112, 264511, 264593, 264760, 265009, 264906, 264769, 264905, 264906, 264907, 264908, 27486282, 59526498, 87165818, 22776900, 264293, 264908, 264910, 265008, 265009, 264636, 264768, 264563
2749	87346307 (5497, 5498)	Novel Protein sim. GBank (g1072426) (U74207) - FIUS (Oryctolagus cuniculus)	UNCLASSIFIED	264488, 52644507, 18100386, 56994075, 264429, 29331825, 29331826, 29331827, 29331828, 264508, 265009, 264910, 264591, 264595, 33657084, 265011, 265019, 18100351, 264288, 264686, 264769, 264689, 55811957, 264693, 27486284, 18100370, 18100374, 264698, 18100385, 264482, 264563
2750	87356344 (5499, 5500)	Novel Protein sim. GBank (g1072426) (U74207) - FIUS (Oryctolagus cuniculus)	UNCLASSIFIED	264488, 52644507, 18100386, 56994075, 264429, 29331825, 29331826, 29331827, 29331828, 264508, 265009, 264910, 264591, 264595, 33657084, 265011, 265019, 18100351, 264288, 264686, 264769, 264689, 55811957, 264693, 27486284, 18100370, 18100374, 264698, 18100385, 264482, 264563
2751	87057485 (5501, 5502)	Novel Protein sim. GBank (g1004185) (AC004634) - OG-2 (Homo sapiens)	Contains protein domain (PF00048) - Homeobox	29331822, 29331824, 265017, 33657023
2752	88062675 (5503, 5504)	Novel Protein sim. GBank (g1004185) (AC004634) - OG-2 (Homo sapiens)	Homeobox domain	

2753	94139872 (5505, 5506)	Novel Protein sim. GBank g 3651648 (AF090301) - neural F-box domain.	Contains protein domain (PF00646)	UNCLASSIFIED	56182575, 56994075, 2271896, 29331822, 29331824, 29331825, 29331826, 295007, 284593, 59012038, 33109954, 18106351, 284288, 59181582, 21906787, 21906788, 265021, 264693, 18108374, 65274791, 264632, 56182323, 22719002, 264563, 264569, 264639, 6971417, 29331827, 35656062, 29331828, 264905, 264908, 264909, 265011, 265017, 265018, 265019, 264288, 21906785, 21906787, 265022, 33657023, 264693, 56182323, 18108382, 22719000
2754	94115513 (5507, 5508)	Novel Protein sim. GBank g 353428 (U13736) - calmodulin-like protein (Plum salinum)	Contains protein domain (PF00036) EF hand	struct	264563
2755	68001472 (5509, 5510)	Novel Protein sim. GBank g 296653 (AC094510) - R03385, 2 Homo sapiens	Zinc finger, C2H2 type	transcript factor	264594
2756	11456038 (5511, 5512)	Novel Protein sim. GBank g 173337 (U30473) - putative src-like adaptor protein; non-catalytic src-like adaptor protein containing SH3 and SH2 domains; homolog of mouse SLAP. Method: conceptual translation supplied by author [Homo sapiens]	Contains protein domain (PF00017) - Src homology domain 2	UNCLASSIFIED	85655542, 264693
2757	85301590 (5513, 5514)	Novel Protein sim. GBank g 3638372 (AF053768) - brain specific contactin-binding protein CBP90 [Rattus norvegicus]	Src homology domain 2	eph	264693
2758	79637846 (5515, 5516)	Novel Protein sim. GBank g 2072200 (U94853) - p40 [Borna disease virus]		struct	264693
2759	91005312 (5517, 5518)	Novel Protein sim. GBank g 2072200 (U94853) - p40 [Borna disease virus]		struct	264693
2760	79624798 (5519, 5520)	Novel Protein sim. GBank g 4091457 (Jen) C48368.11 - (ALG0396) hypodermal protein [Homo sapiens]	Contains protein domain (PF00023) Ack repeat	UNCLASSIFIED	264693
2761	87639597 (5521, 5522)	Novel Protein sim. GBank g 3130150 (AF050182) - PERIOD 3 [Mus musculus]	Contains protein domain (PF00089) - struct	nuclease	264693
2762	87526599 (5523, 5524)	Novel Protein sim. GBank g 3130150 (AF050182) - PERIOD 3 [Mus musculus]	Contains protein domain (PF00089) - struct	struct	264693
2763	87539688 (5525, 5526)	Novel Protein sim. GBank g 3511122 (AF065603) - zinc finger protein [Homo sapiens]	Contains protein domain (PF01332) KRAA box	transcript factor	264693
2764	84305140 (5527, 5528)	Novel Protein sim. GBank g 2906643 (AF045424) - ribitol kinase [Mycobacterium tuberculosis]	Contains protein domain (PF00030) - P-GGT family of carboxylate kinases	struct	264693

2765	94315105 (5529, 5530) Novel Protein sim. GBank g1468872[eml](CAA17888.2) - (A1022018) /prediction=method; /match=(desc: [Drosophila melanogaster])		264485, 65274572, 22778995, 35696288, 22778998, 22778999, 264259, 30331822, 20331824, 35696052, 264508, 264509, 264512, 264906, 264907, 264908, 264909, 264910, 264911, 265009, 264592, 264595, 264758, 58812038, 33109954, 265010, 87168559, 264600, 265018, 264760, 264761, 264762, 264763, 264448, 264764, 264288, 264766, 264767, 264788, 56181552, 21906764, 21906765, 21906768, 35695917, 265021, 264691, 264692, 33657023, 33657109, 264628, 18108370, 264629, 18108374, 59811576, 35696423, 35695955, 264631, 264634, 264635, 264636, 264637, 264638, 294536, 264639, 83373044, 83373045, 22279000, 22279002, 264593, 264595
2766	94315109 (5531, 5532) Novel Protein sim. GBank g16441611[eml](CA046854.1) - (A1388555) hypothetical protein [Cans. familiaris]	Contains protein domain (PF00076) RNA recognition motif (a.k.a. RRM, RBD, or RNP domain)	264486, 65274572, 80432289, 264407, 264908, 264511, 264512, 60433539, 264788, 294685, 294689, 35695917, 265022, 264903, 264628, 63274791, 264635, 264555, 264556, 264557, 264638, 264558, 264559, 83373044, 60432113
2767	80304297 (5533, 5534) Dipodomys deserti muscle - chicken		264412, 263974, 264538
2768	94322238 (5535, 5536) Novel Protein sim. GBank g16441322[eml](CA046721.1) - (A1031427) U187A19.1 (novel protein) [Homo sapiens]		264485, 263994, 264439, 35696286, 22778998, 22778999, 264259, 29331822, 60432289, 29331826, 35696052, 264508, 264509, 264905, 264906, 264907, 29333130, 264908, 264909, 264510, 264511, 264512, 265005, 264910, 265009, 264591, 264592, 264593, 58812038, 264594, 264595, 264757, 264596, 265011, 264596, 264597, 264598, 264691, 18108351, 264764, 264768, 264689, 264690, 264769, 18108357, 264769, 264689, 21906766, 21906768, 21906769, 35695917, 265021, 265022, 33657023, 33657109, 264628, 264629, 18108374, 35696423, 35695955, 264630, 264631, 264632, 264634, 264635, 264636, 264555, 264638, 264639, 83373044, 264563, 264564, 264565, 264566, 264486, 264587

2769	9531168 (5537, 5538) Novel Protein sim. GBank g15(1985jmb)CA8437.5 11- (AL086725) hypothetical protein [Homo sapiens]	ubulin	264488, 56182575, 264259, 29331822, 29331824, 60714117, 29331825, 56182435, 29331826, 264608, 29331830, 56182435, 264592, 33657402, 264448, 264369, 264286, 60170615, 264691, 33657023, 264692, 33657109, 18108374, 5811576, 264634, 264636, 56182323, 63373044, 60432113
2770	87730182 (5539, 5540) Novel Protein sim. GBank g15(70196jmb)CA852157.11- (AL109736) WD repeat protein [Schizosaccharomyces pombe]	Contains protein domain (PF00400) - kinase WD domain, G-beta repeat	18108398, 22778995, 22778996, 22778997, 254031, 265006, 265018, 18108351, 264897, 21906765, 18108364, 264628, 18108374, 264631, 18108385, 18108389, 264599, 29331822, 29331824, 29331825, 264599
2771	86084071 (5541, 5542) Novel Protein sim. GBank g10593433 (AC004129) - Unknown gene product [Homo sapiens]	UNCLASSIFIED	264599, 29331822, 29331824, 60714117, 60432269, 29331827, 264903, 265009, 264592, 55812038, 56274444, 264766, 21906769, 33657109, 263976, 284555, 264638, 264557, 83373044, 264663, 264564, 264486
2772	95357309 (5543, 5544) Novel Protein sim. GBank g14(88553)jmbP_005465.1jmbY C - histone deacetylase 5	Contains protein domain (PF00850) - histone histone deacetylase family	18108374, 264686, 264687, 263976, 56182435, 264689, 5581576, 55810764, 21906766, 35696423, 5581576, 6324791, 56181686, 55811987, 35695855, 264110, 265021, 264112, 265022, 265006, 265008, 264092, 264094, 60431850, 264637, 264891, 264259, 264692, 263981, 264594, 60433356, 56182323, 264693, 264757, 56182111, 55811638, 29331825, 604626, 18108385, 264692, 29331827, 27486261, 29331828, 35695062, 55811386, 264107, 60432113, 265017, 55811150, 18108351, 264681, 264506, 18108370, 264484, 264682, 20281069, 264448, 63712502, 264683, 264764, 264288, 264684, 264766, 263974
2773	94139994 (5545, 5546) Novel Protein sim. GBank g13288888 (AC005253) - R26445_1 [Homo sapiens]	UNCLASSIFIED	22778995, 35696286, 22778996, 22778997, 264529, 29331826, 60432286, 35696052, 264112, 33657402, 21067574, 87188599, 265017, 265018, 264288, 21906766, 21908767, 21900769, 35695917, 265020, 265021, 33657023, 33657109, 18108370, 263976, 35696423, 35695855, 87188518, 12279000, 264482
2774	87819906 (5547, 5548) Novel Protein sim. GBank g165852jmbP343881(753 CAEL - HYPOTHETICAL 70.7 KD PROTEIN FORG8.3 IN CHROMOSOME III	Contains protein domain (PF00380) - ribosomal prot Ribosomal protein S9S16	

2799	38090651 (5597, 5598)	Novel Protein sim. GBank gi 252825 (AC004382) - Unknown gene product (Homo sapiens)	UNCLASSIFIED	22278996, 22278997, 22278998, 22278999, 60432049, 264289, 29331125, 33696002, 29331828, 264909, 60433358, 33657402, 33109954, 87168474, 264448, 52644229, 21906766, 21906767, 21906768, 35959291, 265020, 265021, 265022, 52644150, 35695855, 264634, 60432113, 22277000
2800	88316481 (5699, 5600)	Novel Protein sim. GBank gi 4240301 (p B A V 7429.1) - (AB020713) KIAA0006 protein (Homo sapiens)	glycoprotein	244488, 56994075, 264239, 20281089, 29331823, 29331827, 264905, 58162435, 265008, 21906769, 265009, 265010, 265011, 265012, 264238, 264269, 264686, 60170815, 264691, 264692, 27486265, 264628, 264629, 264636, 264657, 264658, 264659, 87168516, 264654, 264656, 264657, 264658, 265007, 264687
2801	86068814 (5601, 5602)		Contains protein domain (PF00627) - UBA domain	264448, 35695555
2802	86082477 (5603, 5604)	Novel Protein sim. GBank gi 2337665 (AC002464) - organic cation transporter; 50% similarity to JC4884 (P1Dg143892) [Homo sapiens]	transport	264448, 35695555
2803	79577446 (5605, 5606)	Novel Protein sim. GBank gi 45593686 (A023029) - (AC006586) - hypothetical protein [Arabidopsis thaliana]	UNCLASSIFIED	264639
2804	57111131 (5807, 5808)		G-patch domain	264566
2805	87394486 (5609, 5610)		Contains protein domain (PF01585) - G-patch domain	264092, 264259, 29331822, 29331824, 29331827, 29331828, 264908, 265007, 265008, 21906764, 264369, 264286, 264632, 29331823, 29331825, 29331826, 29331827, 29331828, 29331829, 29331830, 29331831, 29331832, 29331834, 29331828, 29331827, 29146496, 265008, 265009, 60433438, 265017, 265018, 265019, 264448, 264428, 21906765, 21906767, 21906769, 29146629, 265022, 52644150, 56182323
2806	87898951 (5611, 5612)	Novel Protein sim. GBank gi 1689731p P44403 (C C P B - CLP B PROTEIN	UNCLASSIFIED	264259

2807	91720702 (5613, 5614)	Novel Protein sim. GBank g144633(jlmbjCAB37931) - (AL031433) dL4S5N24.1 (PUTATIVE novel protein similar to predicted yeast and worm proteins) [Homo sapiens]	UNCLASSIFIED	55844557, 52645150, 53646842, 16108388, 56113257, 22278999, 22278996, 56940075, 35696286, 22278997, 22278998, 22278999, 264259, 52645380, 20331832, 29331824, 29331825, 29331826, 35696002, 29331828, 33655970, 264100, 264105, 264807, 52644045, 60433355, 264584, 60433438, 52948317, 21906754, 33109954, 33657084, 52644296, 87168474, 265011, 87168559, 265017, 265018, 265019, 18108351, 264448, 264763, 264687, 52644229, 21906765, 21906766, 21906767, 21906768, 21906769, 35655917, 265021, 265022, 52644150, 33657023, 52645159, 33657109, 33657182, 27486261, 27486262, 27486264, 27486265, 3565763, 18108370, 18108372, 18108374, 18108376, 263977, 18108377, 35696423, 35696555, 52644332, 85416441, 18108353, 18108357, 87168516, 60432113.
2808	95159111 (5615, 5616)	Novel Protein sim. GBank g14541863(jlmbjCA051071.1) - (AL096857) hypothetical protein [Homo sapiens]	MHC	60432259, 264510, 265010, 265017, 265018, 264681, 264686, 265021, 264680, 22278900, 264565
2809	86083330 (5617, 5618)	Novel Protein sim. GBank g14772561 (AC004002) - similar to dairy d9retn beta heavy chain, 78% Similarity to P2098 (P2D.g118465) [Homo sapiens]	ATPase-associated	18108351
2810	87359032 (5619, 5620)		UNCLASSIFIED	264869, 22278996, 264091, 264259, 29331828, 29146499, 29146529, 29148784, 264683
2811	91735444 (5621, 5622)	Novel Protein sim. GBank g1324453 (AC005189) - match to ESTs H97759 (NIC0111643) and A0485546 (IND.g1628723) [Homo sapiens]	UNCLASSIFIED	264106
2812	86083334 (5623, 5624)	Novel Protein sim. GBank g1424273(jlmbjCA4151.1) - (AB020659) KIAA0832 protein [Homo sapiens]	UNCLASSIFIED	
2813	91716735 (5625, 5626)	Novel Protein sim. GBank g1424273(jlmbjCA4151.1) - (AB020659) KIAA0832 protein [Homo sapiens]	UNCLASSIFIED	56183515, 29331828, 264096, 264096, 60712502, 55811384, 26501, 35696423, 35696555, 264683, 265020, 87168515, 60432113.
2814	90080006 (5627, 5628)	Novel Protein sim. GBank g13440791 (AC005620) - RC3390.1 [Homo sapiens]	Contains protein domain (PF00170) - bZIP transcription factor	65274572
2815	78744321 (5629, 5630)	Novel Protein sim. GBank g1424263(jlmbjCAB4660.1) - (AJ243460) proteoglycan [Leishmania major]	UNCLASSIFIED	264807, 264909
2816	95358229 (5631, 5632)	Novel Protein sim. GBank g1424263(jlmbjCAB4660.1) - (AJ243460) proteoglycan [Leishmania major]	UNCLASSIFIED	264488, 35696286, 29331825, 29331828, 264508, 264509, 264906, 264906, 264907, 264908, 264909, 264511, 264910, 265011, 264760, 264681, 264766, 264769, 264689, 21906765, 264693, 264628, 18108370, 264629, 264631, 264634, 264963, 264584, 264585, 264486

2817	87748542 (5633, 5634)	Novel Protein sim. GBank gll253846 (J56566) - coded for by C. elegans cDNA YK3003.5 coded for by C. elegans cDNA YK3003.3 [Caenorhabditis elegans]			264259, 26331822, 26331827, 264498, 264500, 264505, 265007, 264608, 264609, 56182455, 264510, 265006, 264511, 264512, 264583, 264738, 265010, 264760, 264761, 264764, 264288, 264687, 264768, 5811957, 33685917, 33657109, 263578, 264634, 264638, 264839, 264564, 264565, 264466, 264485, 264567 66712502
2818	86073579 (5635, 5636)	Novel Protein sim. GBank gll44966 (U13149) - possible apoptosis-associated protein [Penicillium citreum]	Contains protein domain (PF00023) - UNCLASSIFIED Ant repeat		
2819	87793267 (5637, 5638)	Novel Protein sim. GBank	UNCLASSIFIED		264909, 264511
2820	87765747 (5639, 5640)	Novel Protein sim. GBank gll4693 (gllg4604-147) [A.15209 - (A515209) CG1-05 protein [Homo sapiens]	Contains protein domain (PF00819) - UNCLASSIFIED UPF004		18100394, 52646365, 52644045, 264112, 264507, 264508, 264509, 264510, 264511, 264512, 264583, 264589, 264688, 264695, 21056769, 33657033, 33657109, 18108370, 18108374, 18108377, 18108385
2821	95320511 (5641, 5642)	Novel Protein sim. GBank gll399144 (gllp02747C1QC, HUMAN - COMPLEMENT C1Q SUBCOMPONENT, C CHAIN PRECURSOR (AB002363) KIAA0365 [Homo sapiens])	Contains protein domain (PF00386) - complement C1q domain		264638
2822	94260221 (5643, 5644)	Novel Protein sim. GBank gll224671 (gllpBAA20820) -	UNCLASSIFIED		264259, 26331822, 60432385, 264905, 60433356, 33657402, 265011, 264760, 21060765, 60170615, 264539
2823	95320513 (5645, 5646)	Novel Protein sim. GBank gll399144 (gllp02747C1QC, HUMAN - COMPLEMENT C1Q SUBCOMPONENT, C CHAIN PRECURSOR	Contains protein domain (PF00386) - complement C1q domain		264768
2824	95320515 (5647, 5648)	Novel Protein sim. GBank gll399144 (gllp02747C1QC, HUMAN - COMPLEMENT C1Q SUBCOMPONENT, C CHAIN PRECURSOR	UNCLASSIFIED		264907
2825	10742170 (5649, 5650)				264760
2826	94111905 (5651, 5652)	Novel Protein sim. GBank gll355965 (gllm062422020) - (AL033503) conserved hypothetical protein [Candida albicans]			52645156, 22778994, 35666286, 22778997, 22778998, 52645080, 26331822, 26331824, 26331825, 26331827, 35666052, 26331828, 35666070, 52644045, 52646317, 33657084, 52644296, 265017, 265018, 265019, 264288, 21906764, 21906765, 21906768, 21906769, 21906769, 35665997, 52644150, 33657109, 33657182, 27486261, 27486262, 33657349, 27486265, 35665763, 35666423, 35668855, 52644332, 18108385, 6718518, 264484

2827	85320519 (5653, 5654)	Novel Protein sim. GBank glt35914sp p02747YC.TOC. HUMAN - COMPLEMENT C1QC SUBCOMPONENT, C CHAIN PRECURSOR	Contains protein domain (PF00336) - complement C1q domain	254488, 263954, 264485, 55274572, 26331822, 86714117, 22331827, 29331828, 264508, 264905, 264906, 264906, 264907, 264908, 264909, 264510, 265006, 264511, 265007, 264512, 265008, 264910, 264591, 60433229, 264592, 264993, 264595, 264596, 2-906754, 265011, 264600, 264601, 264602, 265017, 264604, 264605, 264606, 264706, 264707, 264689, 55811897, 264534, 264590, 264591, 264592, 264593, 264594, 264595, 26498845, 264631, 264634, 264635, 264555, 264636, 264637, 264596, 264638, 264558, 83373044, 18108355, 60432113, 22279002, 264563, 264564, 264565, 264566, 264567, 264486, 18108391
2828	91228612 (5655, 5656)	Novel Protein sim. GBank glt358974 (AF017000) - protein [lysine phosphatase TD14 (Rattus norvegicus)]	Contains protein domain (PF00102) - phosphatase Protein-tyrosine phosphatase	25331822, 35696052, 264104, 204908, 265007, 264591, 265010, 865011, 265019, 264786, 264686, 55811957, 18108370, 18108374, 55810764, 35686423, 55811576, 55182323, 83373044, 87168518
2829	87651244 (5657, 5658)	Novel Protein sim. GBank glH480686glH4AD27734-1 (AF139949) CGI-25 protein [Homo sapiens]		22278996, 22278997, 264091, 264093, 60432049, 264459, 20331822, 26331825, 26331827, 26331828, 264905, 264599, 60712502, 264510, 264511, 264593, 60433498, 21906754, 265011, 264603, 18108355, 264638, 55811957, 2646768, 21906759, 26448634, 52641450, 264693, 33657109, 18108374, 264634, 18108395, 60432113, 22279000, 264565, 264486, 265008, 265019, 264638, 22275002
2830	86087100 (5659, 5660)	Novel Protein sim. GBank gl2468667sp Q61200 NP11_MOUSE - NEUREXOPHILIN 1		265017
2831	87614717 (5661, 5662)		UNCLASSIFIED	22278997, 22278999, 52646317, 264288,
2832	87631800 (5663, 5664)		UNCLASSIFIED	264688, 21906767, 60431528, 264638, 22279000
2833	87612308 (5665, 5666)	Novel Protein sim. GBank glt262815sp Q5CAB45747.1 - (AL080156) hypothetical protein [Homo sapiens]	UNCLASSIFIED	254455, 264556, 264558
2834	86914703 (5667, 5668)	Novel Protein sim. GBank glt2245677sp BAA20772 - (AB002311) KIAA0313 [Homo sapiens]		263972
2835	87715712 (5669, 5670)	Novel Protein sim. GBank glt3389532sp BAA76788.1 - (AB023161) KIAA0944 protein [Homo sapiens]		60422289, 26331828, 265008, 265010, 265011, 26444448, 55811957, 265020,
2836	85724748 (5671, 5672)	Novel Protein sim. GBank glt351568 (U76618) - N-RAP [Mus musculus]	Contains protein domain (PF00412) - transcriptinador LM domain containing proteins	264559, 264112, 265010, 264762, 264764, 263974, 264555, 264558, 264559

2837	87766482 (5673, 5674)	Novel Protein sim. GBank g15420387[emb CA846679.1]- (AJ243459) protozoophagocytosis [Leishmania major]	UNCLASSIFIED	18106394, 22278697, 22278698, 29331822, 284907, 264909, 285006, 285007, 265016, 255019, 264682, 264288, 21906766, 21906767, 55811957, 35695917, 18108374, 56182223, 22279000, 22279002, 56182223, 18108394, 18108397, 264459, 29331826, 265007, 265019, 264448, 18108368, 20281149, 264565, 264567
2838	87775392 (5675, 5676)	Novel Protein sim. GBank g1073378 (U-1263) - core protein [Hepatitis C virus]	UNCLASSIFIED	254555
2839	85799317 (5877, 5878)	Novel Protein sim. GBank g1157515 (JK-4889) - thrombospondin-related anonymous protein [Plasmodium gallinaceum]	UNCLASSIFIED	254509, 264511, 265011, 264288, 264769, 265020, 264634, 264638, 264596
2840	87774855 (5679, 5680)	Novel Protein sim. GBank g1224605[gi AA20790] - (AB022330) KIAA0332 [Homo sapiens]	Contains protein domain (PF00076) RNA recognition motif (a.k.a. RRM, RBD, or RNP domain)	5618235, 35696052, 264907, 264908, 264909, 264910, 264593, 264595, 264766, 264911, 264912, 264913, 264914, 18108370, 35695855, 264631, 264559, 264663, 264597
2841	86582568 (5681, 5682)	Novel Protein sim. GBank g1224605[gi AA20790] - (AB022330) KIAA0332 [Homo sapiens]	UNCLASSIFIED	264600
2842	80030086 (5683, 5684)	Novel Protein sim. GBank g15578957[emb CA81350.1] - (AL050300) d417587.1 (novel KIAA0616 and KIAA0323 LINE protein) [Homo sapiens]	UNCLASSIFIED	254906, 264907, 264908, 264909, 264910, 204784, 35695855, 83373044, 18108385
2843	91012494 (5685, 5686)	Novel Protein sim. GBank g155123[gi Q0878]FBL_C_MOUSE - FIBULIN-1, ISOFORM C PRECURSOR [HOUSE-MOUSE - PROTEIN 80] (BM-90)	Contains protein domain (PF00038) - EGF-like domain	264685
2844	56731154 (5687, 5688)	Novel Protein sim. GBank g15420387[emb CA846679.1] - (AJ243459) protozoophagocytosis [Leishmania major]	homeobox	29146498, 87768474, 264686, 35696423, 83373044, 264564
2845	94321719 (5689, 5690)	Novel Protein sim. GBank g15420387[emb CA846679.1] - (AJ243459) protozoophagocytosis [Leishmania major]	UNCLASSIFIED	29331830, 264909, 265008, 265011, 87169593, 264629, 264556
2846	88318513 (5691, 5692)	Novel Protein sim. GBank g1530663[gi AA1995]TAC00623 - (AC006233) Unknown protein [Neurospora crassa]	UNCLASSIFIED	264608, 264766, 264769, 264629, 264637, 264586
2847	81811757 (5693, 5694)	Novel Protein sim. GBank g1539676 (AC053360) - R01.160.1 [Homo sapiens]	UNCLASSIFIED	264480, 264239, 264508, 264905, 264907, 264510, 265007, 265008, 264591, 264982, 264593, 264594, 264595, 5812038, 264909, 265011, 264604, 264763, 264764, 264765, 264766, 264686, 264628, 264629, 264555, 264636, 264556, 264557, 264638, 264559, 264559, 264563, 264566, 264567
2848	87612943 (5695, 5696)	Novel Protein sim. GBank g1532651[emb CA846747.1] - (AL080165) hypothetical protein [Homo sapiens]	UNCLASSIFIED	29331822, 35696052, 264509, 264906, 265007, 264594, 265016, 264288, 263972, 35696423, 18108394, 56556486, 18108390, 35696286, 265008, 265009, 265018, 204288, 36959917, 264693, 18108374, 35695855, 22279000
2849	88084283 (5697, 5698)	Novel Protein sim. GBank g1342218 (AC004131) - Unknown gene product [Homo sapiens]	UNCLASSIFIED	
2850	87623656 (5699, 5700)		UNCLASSIFIED	

2851	0782354e (5701, 5702)	Novel Protein sim. GBank g1321619g[AAO15788.1] - (AF05 0098) seven transmembrane domain orphan receptor [Mus musculus]	UNCLASSIFIED	264900, 264907, 5618345, 264758, 55911386, 265010, 18108351, 264448, 264569, 21908785, 254691, 264692, 264693, 35659555, 264556, 18108385, 264587, 264451
2852	86397023 (5703, 5704)	Novel Protein sim. GBank g11825729 (U88305) - similar to anophila membrane protein PATCHED SP-P1 (5502) [Ptd.g12845] [Caenorhabditis elegans]	UNCLASSIFIED	56182575, 558117150, 264650, 27486262, 27486265, 264632, 56182323, 56584848, 27279002
2853	87184630 (5705, 5706)	Novel Protein sim. GBank g12702347 (AF027503) - putative membrane-associated guanylate kinase 1 [Mus musculus]	UNCLASSIFIED	35692386, 29331824, 29331826, 26331828, 264508, 264768, 264693, 22279002, 264482
2854	88083357 (5707, 5708)	Novel Protein sim. GBank g12785825 (AC004021) - leucine protein, ring canal component involved in cytoplasmic bridges, 77% Similarity to A45773 (P1Cg107908) [Homo sapiens]	UNCLASSIFIED	22278894, 29331822, 29331824, 29331825, 264508, 264908, 265009, 33108954, 265010, 265019, 264448, 21906765, 265020, 264880, 27486265, 83373044, 22279000, 22279002, 264482
2855	94723856 (5709, 5710)	Novel Protein sim. GBank g11504040[b]BA132191 - (D89863) similar to Dmelanogaster peroxidasin(U11052) [Homo sapiens]	UNCLASSIFIED	21906766, 22278697, 255022, 29331822, 29331826, 27486265, 265007, 265009, 265017, 264482, 264563, 18108351, 22278895, 35696286, 29331824, 29331825, 35696052, 264103, 264108, 56182335, 21906765, 21906768, 255020, 18108388, 35655763, 22279002, 264453
2856	88630350 (5711, 5712)	Novel Protein sim. GBank g1264583 (AC005189) - match to EST's H97756 (ND-g1118643) and A4085546 (ND-g1528723) [Homo sapiens]	UNCLASSIFIED	264360, 264887, 22278895, 22278896, 22278897, 22278898, 264459, 29331826, 29331827, 29331828, 264905, 264906, 264907, 264908, 3264045, 264909, 264515, 265010, 265011, 265012, 265013, 265014, 264909, 264509, 264590, 264910, 265015, 264585, 264758, 21906754, 265010, 265011, 264694, 265018, 264760, 15108351, 264763, 264682, 264764, 264765, 264288, 264359, 264685, 264766, 264768, 18108357, 264769, 21906766, 21906767, 265021, 264534, 60170615, 264691, 264692, 18108370, 264629, 18108374, 264631, 264636, 263981, 18108381, 264558, 18108385, 22279002, 264584, 264586, 264481, 264457
2858	87434748 (5715, 5716)	Novel Protein sim. GBank g1423585[pp33227ME18, HUMAN - DNA-BINDING PROTEIN MEL-18 (ZINC FINGER PROTEIN 144)]	UNCLASSIFIED	60424179, 65274572, 29331828, 265005, 264511, 264758, 265011, 21906767, 21906769, 55811957, 255021, 265021, 56182323
2859	00637975 (5717, 5718)	Novel Protein sim. GBank g1423585[pp33227ME18, HUMAN - DNA-BINDING PROTEIN MEL-18 (ZINC FINGER PROTEIN 144)]	UNCLASSIFIED	60424179, 65274572, 29331828, 265005, 264511, 264758, 265011, 21906767, 21906769, 55811957, 255021, 265021, 56182323

2863	95303283 (5735, 5736)	Novel Protein sim. GBank gi 1292686 emb CA63923 - (X94232) I-Cell activation protein [Homo sapiens]	18108362, 264488, 22278994, 22278997, 22278998, 22278998, 60432049, 264259, 29331822, 29331824, 29331825, 80432289, 29331826, 29331827, 35686052, 29331828, 264508, 52644045, 264428, 265006, 265007, 265008, 265009, 264591, 60432223, 264593, 60433355, 264595, 21906754, 265017, 265019, 264682, 264369, 21906705, 21906766, 21906767, 21906768, 265021, 265022, 264891, 33657182, 18103388, 27486291, 27486292, 27486294, 27486295, 18103370, 18103374, 35686053, 35686055, 18103376, 27486293, 27486296, 27486297, 22279000, 22279002, 264482, 264583, 264584, 264587, 264487
2869	88094412 (5737, 5739)		UNCLASSIFIED
2870	84404574 (5739, 5740)		UNCLASSIFIED
2871	88318621 (5741, 5742)	Novel Protein sim. GBank gi 5306263 gb AAD41995.1 AC00623 unknown protein [Arabidopsis thaliana]	264805, 264908, 264764, 21905759, 264634, 264259, 29331822, 60432289, 29331827, 264907, 265008, 265017, 265018, 264682, 264764, 18108354, 865021, 27486285, 264629, 18108357, 264597
2872	95312197 (5743, 5744)	Novel Protein sim. GBank gi 122053 p I B39065 - proline-rich protein 15 - rat	263981
2873	88094252 (5745, 5746)		UNCLASSIFIED
			264488, 18108374, 264766, 264687, 264688, 264689, 35996423, 35996296, 35995917, 264510, 264511, 265007, 264512, 265008, 264910, 264834, 264834, 264635, 264355, 264382, 264289, 264588, 60433438, 264389, 264390, 264391, 264392, 264393, 264394, 264395, 264396, 264397, 264398, 264399, 264400, 264401, 60432113, 264608, 264563, 264482, 264509, 264605, 264908, 264654, 18108351, 264763, 18108370, 264907, 264566, 264908, 264764, 264288, 264587, 264908, 264486, 264766, 18108391

2874	84313546 (5747, 5748)	Novel Protein sim. GBank g1321354 (AC004005) - unknown protein [Arabidopsis thaliana]	UNCLASSIFIED	52645507, 52645156, 52645955, 56182575, 22278994, 22278985, 59940775, 35969286, 60432048, 264259, 52645980, 29331822, 56182181, 29331824, 60432059, 29331825, 88714117, 29331826, 29331827, 35969002, 29331828, 33656970, 66712502, 52644045, 265007, 265009, 60433356, 264758, 55812038, 18108346, 52646317, 33108954, 33657084, 265017, 264804, 285018, 285019, 264682, 264369, 264268, 264668, 52642229, 21906766, 21906767, 21906768, 55811957, 295020, 265021, 285022, 52644150, 33657033, 52645129, 18108374, 18108376, 35969443, 56182323, 18108387, 87168518, 60432113, 22279000, 22279002, 264583, 264585
2875	88083726 (5748, 5750)	Novel Protein sim. GBank g12781386 (AC004010) - similar to Leucine-rich transmembrane proteins; 44% similarity to U42767 (PIDg1736919) [Homo sapiens]	Contains protein domain (PF00560) - glycoprotein Leucine Rich Repeat	22278986, 22278987, 22278989, 29331826, 29331828, 29146496, 66712502, 265008, 265017, 18108351, 264685, 264689, 21906767, 18108376, 18108377, 55811576, 60170384, 22279000, 264487
2876	88090854 (5751, 5752)	Novel Protein sim. GBank g12759530 (AC004449) - R33663_2 [Homo sapiens]	Contains protein domain (PF00167) - Igf Fibroblast growth factor	52645385, 55274572, 56182575, 22278997, 22278998, 22278999, 60432048, 52645080, 60424289, 60432289, 29331827, 35969052, 29331828, 66712502, 52644045, 56182435, 60433356, 33657402, 33657084, 265019, 55811150, 264448, 264369, 21906766, 21906768, 21906769, 265020, 33657023, 33657109, 33657162, 27486262, 264629, 60431828, 55811576, 52644332, 56182323, 55811577, 52644333, 52644334, 52644335, 60432113, 22279000, 22279002, 264583, 264585
2877	94747029 (5753, 5754)	Novel Protein sim. GBank g1704208(jmb)(CAB41646.1) - (AL035419) (J1100H13.1 (putative novel protein) [Homo sapiens])	UNCLASSIFIED	264468, 264259, 29331822, 29331826, 264905, 264509, 264907, 264809, 264510, 265006, 264511, 264512, 33657402, 264758, 21906754, 18108351, 264681, 264682, 264288, 264684, 264685, 264768, 264689, 21906769, 264690, 33657023, 264693, 18108384, 33657109, 264629, 18108374, 264630, 264632, 264556, 264637, 264639, 264558, 18108385, 18108387, 264593, 264584, 264585, 264586, 264486, 264587
2878	88095309 (5755, 5756)	Novel Protein sim. GBank g1387675(jmb)(CAB30367) - (Z8107) predicted using GeneSifter. Similarity to Yeast protein B248 (TR-G58753) [Carnochaditis elegans]	UNCLASSIFIED	

2879	87869122 (5757, 5758)	Novel Protein sim. GBank gll4855145gplAD32192.11 - (AF127374) unknown [Streptomyces lavendulae]	Contains protein domain (PF00315) - UacH-DNA glycosylase	UNCLASSIFIED	18108359, 264250, 264905, 18108370, 264829, 264908, 264909, 18108374, 18108377, 265006, 264910, 264637, 60170394, 264559, 265017, 264584, 264565, 264587, 264684, 264339
2880	94451435 (5759, 5760)	Novel Protein sim. GBank gll4680703gplAD27741.1 (AF13296) C01-32 protein [Homo sapiens]			264488, 52640365, 52646842, 2278594, 35696266, 22778959, 22778999, 264299, 28331822, 25331824, 29331825, 29331826, 28331827, 35696052, 29331828, 264107, 264908, 264509, 264905, 264906, 264907, 264908, 52644045, 264510, 265006, 264511, 265007, 264512, 265005, 264910, 264584, 21900754, 52646317, 52644236, 264905, 264906, 264907, 264908, 264784, 264288, 264786, 264788, 264687, 264789, 21905786, 21905793, 35695917, 265021, 264680, 264682, 33657023, 52645129, 33657109, 33657182, 27486262, 33657349, 264629, 18108374, 35695855, 264634, 264635, 264636, 264637, 264638, 264557, 52644332, 264558, 264559, 83373044, 264404, 22776000, 264583, 264483, 264567, 264486
2881	87656539 (5761, 5762)	Novel Protein sim. GBank gll7333571 (U23452) - No definition line found [Caenorhabditis elegans]		UNCLASSIFIED	22278998, 29331822, 52640345, 21906765, 264639, 60432113
2882	8714387 (5763, 5764)	Novel Protein sim. GBank gll118112 (U41559) - No definition line found [Caenorhabditis elegans]			264488, 22278996, 22278999, 28331822, 29331826, 264908, 60170331, 60433396, 5801208, 264681, 264682, 264686, 264687, 264688, 21905786, 21905788, 264683, 264687, 18108374, 58011576, 5818323, 22278002, 264586
2883	95362875 (5765, 5766)	Novel Protein sim. GBank gll468008gplAD31087.1 (AF10693) - (AF10693)4) albumin D receptor-interacting protein [Homo sapiens]	Contains protein domain (PF00400) - VID domain, C-beta repeat	UNCLASSIFIED	264112, 264682
2884	87784643 (5767, 5768)	Novel Protein sim. GBank gll224657 (gll2BAA20832) - (AB02376) KIAA0378 [Homo sapiens]		UNCLASSIFIED	265018, 264634
2885	83006306 (5769, 5770)	Novel Protein sim. GBank gll255889 (U53344) - T07H6.5 gene product [Caenorhabditis elegans]	Contains protein domain (PF00084) - complement receptor Sushi domain (SCR repeat)		264886, 264693
2886	81737823 (5771, 5772)	Novel Protein sim. GBank gll255889 (U53344) - T07H6.5 gene product [Caenorhabditis elegans]			60432049, 264239, 29331828, 264908, 264511, 264595, 60433438, 264596, 285017, 264905, 263969, 263972, 264555, 83373044, 87186518, 264566
2887	91227860 (5773, 5774)	Novel Protein sim. GBank gll388232 (gll3BAA34521.1) - (AB018343) KIAA0801 protein [Homo sapiens]		UNCLASSIFIED	18108351, 264686, 264625, 264031, 264899, 83373044, 264482
2888	85105816 (5775, 5776)	Novel Protein sim. GBank gll468001gplAD_003447.1 (p2472) - zinc finger protein 205 zinc finger, C2H2 type	Contains protein domain (PF00086) - transcription factor zinc finger, C2H2 type		264488, 264259, 29331828, 264508, 264506, 264592, 264593, 264786, 264789, 18108374, 83373044, 264486

2883	87636582 (5777, 5778)			UNCLASSIFIED	5694075, 26331824, 265009, 264760, 18108354, 264286
2883	87636583 (5778, 5780)	Novel Protein sim. GBank g 854065 ame CAA58337 - (X83413) U88 [Human herpesvirus B]		UNCLASSIFIED	264581, 264786
2881	88069428 (5781, 5782)	Novel Protein sim. GBank g 387750 emb CA801508 - (Z78054) predicted using Gene-finder, similar to collagen; cDNA EST EMBL-D65865 comes from this gene; cDNA EST EMBL-D69451 comes from this gene; cDNA EST EMBL-D66028 comes from this gene; cDNA EST EMBL-D66552 comes from this gene...		UNCLASSIFIED	264591, 264585, 264369, 264685, 264683, 264628, 264583, 264568
2882	85419745 (5783, 5784)	Novel Protein sim. GBank g 4829759 ga AAD4140.1 AF15190 - (AF151903) CG-143 protein [Homo sapiens]		UNCLASSIFIED	68274572, 35666626, 26331828, 264110, 265009, 60433438, 265018, 265018, 18108351, 264288, 21900765, 21906766, 21906763, 55811957, 264690, 55274620, 263567, 35695835
2883	87798014 (5785, 5786)			UNCLASSIFIED	264581, 265010, 264682, 18108370, 264535, 264535
2884	87755985 (5787, 5788)	Novel Protein sim. GBank g 1568015 gb CA046135.1 - (AF080171) zinc finger protein ZNF232 [Homo sapiens]	Contains protein domain (PF00056) - Zinc finger, C2H2 type	UNCLASSIFIED	264259, 265008, 60433438, 57842398, 265011, 264389, 35695817, 18108391, 18108392, 18108398
2885	86038778 (5789, 5790)	Novel Protein sim. GBank g 3924708 emb CAA84848 - (Z35597) Weak similarity with sea squirt nidogen precursor protein (blastp score 71); cDNA EST EMBL-T02069 comes from this gene; cDNA EST EMBL-D76135 comes from this gene; cDNA EST EMBL-D73147 comes from this gene; cDNA EST EMB...	Contains protein domain (PF01437) - Plein repeat		26331824, 265007, 264767, 264636, 264583
2886	87732122 (5791, 5792)	Novel Protein sim. GBank g 4485549 ref NP_005456.1 pPKG - protein kinase B gamma	Contains protein domain (PF00069) - Kinase Eukaryotic protein kinase domain		18108392, 18108394, 18108398, 22278998, 264259, 26331822, 26331824, 26146489, 264908, 264908, 265007, 265008, 265018, 265019, 264389, 264685, 264688, 21906766, 265021, 264683, 33837182, 264635, 18108391, 18108392, 18108398
2887	86413007 (5793, 5794)	Novel Protein sim. GBank g 4485549 ref NP_005456.1 pPKG - protein kinase B gamma	UNCLASSIFIED		60433438, 18108370, 22278998, 35696246, 22278997, 22278998, 60433438, 264259, 26331824, 26331828, 60433438, 60432289, 26331826, 26531828, 35696032, 264638, 56182435, 265009, 264910, 60170831, 80431735, 60433356, 60433438, 65274444, 55811396, 265018, 18108351, 264448, 264288, 264687, 21906765, 21906766, 21906767, 21906769, 55811957, 35695817, 264534, 33657023, 33657109, 35695783, 264628, 264629, 60431528, 18108374, 55810764, 55811576, 35696423, 35696385, 264955, 56182323, 18108395, 264404, 22279300, 22279302, 264396
2888	87750340 (5795, 5796)			UNCLASSIFIED	22278995, 264604, 18108395, 264566

2899	00357670 (5797, 5799)	Novel Protein sim. GBank gij1581470jmi[CAB40137.1] - (Y19483) SLC7A8 protein [Homo sapiens]	UNCLASSIFIED	264764, 21906764, 264692 65274572, 66182575, 35656288, 60433049, 264256, 29331824, 66714117, 29331836, 35690492, 29331828, 66712502, 56182435, 265006, 265007, 265008, 265009, 60433356, 264758, 265018, 264764, 264765, 264766, 264768, 21906764, 21906766, 21906769, 265020, 264692, 264693, 32833986, 264631, 83373044, 60432113
2901	07444731 (5801, 5803)	Novel Protein sim. GBank gij159227eRNP_004614.1pTTC4 - heratiscapipide repeat domain 4	phosphatase	22278995, 22278997, 22278999, 60432049, 29331822, 29331824, 29331825, 29331827, 29331828, 35656760, 244910, 265009, 265010, 265011, 265012, 265013, 265014, 265015, 21906764, 19367063, 21906766, 21906767, 21906769, 33657023, 264693, 33657109, 33657144, 35656423, 39669855, 263981, 56182323, 22279002 264683, 264681
2902	05745271 (5803, 5804)	Novel Protein sim. GBank gij2414615jmi[CAB16394] - (Z99259) hypothetical protein [Schistosoma haematophyllum]		264687, 22278994, 264255, 29331826, 29331828, 264905, 52644045, 56182435, 264511, 265017, 265018, 18108351, 264448, 264683, 264769, 264689, 39695917, 52644150, 87166515, 60432113, 22279002
2903	07606733 (5805, 5806)	Novel Protein sim. GBank gij1070318pji[S62241 - XLCl2 protein - African clawed frog]		52644150, 87166515, 60432113, 22279002
2904	06158077 (5807, 5808)	Novel Protein sim. GBank gij593823pji[AD15885.1AF14367 - (AF143676) multipass transmembrane protein with 10 transmembrane domains]	UNCLASSIFIED	52644150, 87166515, 60432113, 22279002
2905	84449326 (5809, 5810)	Novel Protein sim. GBank gij728837spjP3519AJLU7_HUMAN - IIII ALL SUBFAMILY SQ WARNING ENTRY IIII	oncogene	265011, 265017, 264683, 264769, 35695917, 265006, 265007, 18108374, 35656855, 265009, 264692, 264693, 32833986, 18108387
2906	95341051 (5811, 5812)	Novel Protein sim. GBank gij489259pji[AD27831.1AF12185 - (AF12185) sorting nexin 8 [Homo sapiens]]	UNCLASSIFIED	22278995, 35696286, 22278998, 264259, 60432289, 29331828, 29331830, 66711502, 265009, 60170831, 33109554, 264448, 264683, 264288, 264689, 21906766, 21906767, 21906769, 55811957, 35695917, 265022, 52644150, 264691, 33657023, 264692, 264693, 35695855, 60432113, 264696

2807	91211383 (5813, 5814)	Novel Protein sim. GBank gll707079 (U0451) - contains strong similarity to a DNA-like domain (PS-P50636) (Ctenorhabdalis elegans)	Contains protein domain (PF00226) - eph DnaJ domain	52644507, 56182575, 56181886, 56227895, 56994075, 35690286, 60432049, 56182181, 35690502, 80431735, 284595, 55812038, 21906754, 55811386, 255019, 264682, 264369, 56181562, 21906766, 55811987, 35695917, 265020, 255021, 33857023, 33657109, 60431528, 55811576, 35699423, 35695855, 264638, 22279000
2808	80414246 (5815, 5816)	Novel Protein sim. GBank gll873917 (AC020561) - putative ATP-dependent RNA helicase (Arabidopsis thaliana)	helicase	255009, 33109994, 18100351, 254766, 255021, 254891, 264892, 18100374, 254556, 254538, 254537, 254559
2809	87420224 (5817, 5818)		eph	254538, 87168474, 260018, 18100365, 22278995, 264509, 364512, 255007
2910	86601075 (5819, 5820)	Novel Protein sim. GBank gll453333jmmj(CA837483.1) - (AL0354539) putative protein (Arabidopsis thaliana)		35657402, 265017, 264385, 265022, 18100365, 264628
2911	94216615 (5821, 5822)	Novel Protein sim. GBank gll4483187jmmj(CA838415.1) - (AL031569) gll1163J1.3 (mouse protein similar to mouse B99) (Homo sapiens)	glucanase	52645355, 18100397, 22278995, 22278997, 22278998, 22278999, 29331824, 29331825, 52644045, 265006, 255014, 264448, 21906765, 21900766, 21906767, 21900768, 21900769, 55811957, 255021, 18100370, 18100372, 18100374, 22279000
2912	87731803 (5823, 5824)	Novel Protein sim. GBank gll4292637jggAAX-4079 (JAF 15184 - (AF151842) CGI-84 protein (Homo sapiens)	Contains protein domain (PF-00004) - invulcin repeat	52645156, 264092, 60432049, 264259, 52645080, 29331821, 29331825, 66712502, 33109994, 294780, 264683, 264588, 255021, 264693, 18100365, 253576, 254404
2913	87713823 (5825, 5826)	Novel Protein sim. GBank gll85466jmmj(CA458337) - (X03413) DB (Human herpesvirus 8)	Contains protein domain (PF00096) - UNCLASSIFIED Zinc finger, C2H2 type	52644507, 35454136, 52646842, 56182375, 255019, 264682, 255020, 29331824, 29331825, 29331826, 264683, 29331827, 35690285, 29331828, 264683, 52644045, 56182435, 55812038, 52646317, 21906754, 52944296, 87168474, 265017, 255018, 255019, 18100351, 264682, 264686, 264589, 21905755, 21905766, 21906767, 21906768, 35695917, 255020, 50844150, 27486261, 27486262, 27486265, 35695763, 55811576, 35695855, 52644332, 22279000, 22279002, 264563
2914	87797300 (5827, 5828)			264567

2915	85031672 (5329, 5332)	Novel Protein sim. GBank glj5174485[enP_006030_1]p[IAA - endocytic receptor (macrophage mannose receptor family)]	Contains protein domain (PF00059) - eph Lectin C-type domain	264569, 264488, 264987, 264758, 21906786, 5064642, 21906757, 21906788, 56182575, 29146529, 35695917, 22278996, 22278997, 22278998, 265021, 22278999, 52644150, 264691, 264269, 60432040, 264092, 52845129, 33657109, 33657182, 29331827, 27486261, 35690052, 29331828, 27462622, 27486264, 27486265, 33657349, 29146498, 29146499, 264906, 264907, 18108370, 264908, 18108372, 52844545, 18108374, 56182435, 35695955, 264112, 264510, 265006, 60432228, 264593, 60433356, 56182423, 18108382, 35932083, 18108385, 35932084, 18108386, 35932087, 18108388, 35932089, 18108390, 35932091, 18108392, 265010, 265011, 60432113, 265017, 265018, 22279000, 265019, 55811160, 264681, 18108351, 264753, 264448, 264683, 264556, 18108354, 264359, 264285, 264709
2916	85337790 (5831, 5832)	Novel Protein sim. GBank glj5104851[oljBAA80165.1] - (A-P000061) 305aa long hypothetical dTDP-4- dehydrothymine reductase [Aeropyrum pernix]	dehydrogenase	52645156, 65274572, 22278994, 22278995, 35696286, 22278996, 22278997, 22278998, 22278999, 264259, 29331822, 29331824, 29331825, 69714117, 60432289, 29331826, 29331827, 29331828, 33656970, 264599, 264906, 29331830, 52644045, 264909, 56182435, 60700831, 264592, 264593, 33657402, 60433359, 52646317, 21907754, 35109954, 33657084, 52644286, 69658542, 265011, 29331829, 265018, 18108351, 26468, 264218, 23691523, 21906786, 21906787, 21906788, 21906789, 55811957, 35695917, 265021, 265022, 52644150, 33657023, 33657109, 33657182, 27486261, 27486262, 27486284, 35695763, 18108376, 56811576, 35698423, 65274791, 35695955, 52644332, 264457, 264638, 56182323, 18108387, 67168518, 22278002, 264482
2917	87454546 (5833, 5834)	Novel Protein sim. GBank glj3169065[emjCA10250.1] - (A1023704) putative translocation elongation factor-Tu family [Schistosoma carmichaeli]	UNCLASSIFIED	60433438, 264602, 264682, 6718518, 60432113
2918	65960520 (5835, 5836)	Novel Protein sim. GBank glj339218[piljS30038 - hypothetical protein YK.201c - yeast (Saccharomyces cerevisiae)]		264638
2919	87441497 (5837, 5838)	Novel Protein sim. GBank glj5564935 (AF030001) - unknown [Mus musculus]		66714117, 6671502, 263981

2820	8776932 (5833, 5840)			<p>3569208, 2277697, 264569, 5564590, 26331874, 26331826, 26331827, 264682, 264909, 56182435, 264511, 261768, 37109954, 21906754, 5264286, 265010, 265011, 264601, 265017, 265019, 264681, 264887, 21906767, 265021, 52641150, 264690, 264691, 264692, 264693, 33657109, 33657182, 27486262, 27486284, 27486285, 35696423, 35695855, 264632, 264636, 264637, 264638, 56182323, 60170394, 18103385, 87768516, 60432113</p> <p>35696206, 22776997, 264591, 264092, 264094, 264299, 26331822, 29331824, 29331826, 26331827, 35696052, 2914698, 264104, 264105, 264107, 264508, 264110, 264509, 264511, 60432113, 60432114, 8718434, 265011, 8108351, 263281, 21906765, 21906766, 21906767, 21906768, 21906769, 35695917, 265021, 263974, 18108374, 263978, 263977, 18108376, 264555, 263081, 55626486, 87168518, 22279000, 22279002</p> <p>264605, 264909, 264511, 265006, 265008, 264593, 33657402, 60174339, 18108331, 264783, 21906785, 29148627, 35695917, 264692, 264629, 263976, 55611576, 35695855, 264555, 264558, 56182323, 60170394, 22279000, 264486</p> <p>264488, 18108397, 22278995, 22278996, 22278997, 22278998, 22278999, 29331825, 29331826, 29331827, 29331830, 264511, 265009, 33657402, 265011, 265017, 265018, 264683, 18108394, 21906765, 21906767, 264684, 264685, 264686, 264687, 264688, 264689, 264690, 264691, 264692, 33657109, 33657182, 18108376, 264683, 264636, 18108385, 18108397, 22279000, 264553, 264556</p> <p>265017, 264628, 20281125, 264556</p>
2821	8763982 (5841, 5842)	Novel Protein sim. GBank g1458003ig145804202 [J183194 - (U83194) TRAF4-associated factor 2 [Homo sapiens]]	Contains protein domain (PF00787) - PX domain	<p>56182575, 35696286, 264559, 60432289, 29331827, 264508, 26444045, 264910, 264591, 60432229, 55812038, 21906754, 264681, 264448, 264683, 264286, 264685, 52644229, 264689, 21906785, 21906786, 21906788, 21906789, 265021, 265022, 60170615, 264692, 33657023, 264693, 33657109, 35695923, 65274791, 56182323</p>
2822	87748762 (5843, 5844)	Novel Protein sim. GBank g14580514ig145806779 [1 - (AB023152) KIAA0935 protein [Homo sapiens]]	Contains protein domain (PF01074) - kinase Glycosyl hydrolases family 38	
2823	85337789 (5845, 5846)	Novel Protein sim. GBank g14333268ig14332689 [C9A42888.21 - (283844) G3TE 16.4 (similar to mouse p116tp protein) [Homo sapiens]]	Contains protein domain (PF00109) - struct PH domain	
2824	87731987 (5847, 5848)	Novel Protein sim. GBank g12133095ig12132254 - ribosomal protein L36, mitochondrial - yeast (Saccharomyces cerevisiae)	Contains protein domain (PF00444) - ribosomal Ribosomal protein L36	
2825	95090120 (5849, 5850)	Novel Protein sim. GBank g1268898ig1268899 [CAB11718 - (298980) actin associated protein [Schizosaccharomyces pombe]]	UNCLASSIFIED	

2528	95341003 (5851, 5852)	Novel Protein sm. GBank g12785433 (AF098160) - similar to Arabidopsis thaliana male sterility protein 2 (SW-CD889) [Caenorhabditis elegans]			29331828, 265011, 264768, 264689 264764, 264388, 264630, 264637
2627	90408018 (5853, 5854)	hydroxyproline-rich glycoprotein - perennial leontine			UNCLASSIFIED
2628	20452179 (5855, 5856)	Novel Protein sm. GBank g12413320 (emb) (CAA-26915) - (A-000219) CMP-N-acetylneuraminic acid synthetase [Mus musculus]			UNCLASSIFIED
2629	91622920 (5857, 5858)				264559, 264489, 22278994, 35696286, 22278998, 22278999, 22278999, 264094, 264259, 50645090, 29331822, 29331824, 56714117, 29331825, 29331826, 29331827, 35696052, 33659970, 264109, 29331830, 52644045, 265009, 33109954, 52644286, 27166559, 264780, 264762, 264448, 264764, 264288, 264766, 264768, 21906765, 21906766, 21906768, 21906769, 35695917, 264891, 35657023, 264693, 33657109, 31803174, 263876, 35696423, 35695855, 264841, 22279600, 22279602, 264567, 264486
2930	95302755 (5859, 5860)				UNCLASSIFIED
2931	94312693 (5861, 5862)	Novel Protein sm. GBank g12785433 (AF098160) - similar to Arabidopsis thaliana male sterility protein 2 (SW-CD889) [Caenorhabditis elegans]			22278998, 22278999, 264295, 29331825, 80432289, 29331828, 264905, 52844045, 56162435, 265009, 60170331, 264592, 60432223, 60433356, 87168474, 265010, 265011, 265017, 265018, 265019, 264762, 264448, 264683, 264288, 264766, 21906765, 21906769, 35695917, 60170615, 33657023, 33657109, 264628, 18108370, 18108372, 35696423, 35695855, 264556, 56182323, 60432113, 264567
2932	94312693 (5861, 5862)	Novel Protein sm. GBank g12785433 (AF098160) - similar to Arabidopsis thaliana male sterility protein 2 (SW-CD889) [Caenorhabditis elegans]			UNCLASSIFIED
2933	91720776 (5865, 5866)	Novel Protein sm. GBank g1378056 (AF017777) - helicase [Drosophila melanogaster]			helicase
2934	91720776 (5865, 5866)	Novel Protein sm. GBank g1378056 (AF017777) - helicase [Drosophila melanogaster]			264506, 264307 264488, 18108392, 56182576, 22278999, 264091, 264259, 29331825, 60432269, 29331827, 264508, 52644045, 56182435, 265007, 265009, 264592, 60433356, 60433438, 21906754, 265017, 264682, 264288, 52644229, 21906765, 21906766, 21906768, 21906769, 255022, 52644190, 33657023, 33657109, 27466265, 264635, 264636, 60170394, 56182323, 18108395, 60432113, 264565, 264566, 264567

2934	86576025 (5877, 5869)				22278997, 22278998, 29331824, 33657402, 264691, 27486262, 264628, 87168518, 22279000
2935	86410579 (5869, 5870)			UNCLASSIFIED	56182575, 22278995, 69043336, 33657402, 264788, 33100954, 21900754, 265018, 265019, 264448, 264769, 21908764, 21008765, 265021, 264692, 33657023, 33657109, 33657349, 55810784, 22279000, 22278997, 29331827, 29331828, 265009, 265017, 264603, 265020, 56115176, 6108397, 60432113, 264543, 264691, 264788, 265018, 265019, 29331827, 29331828, 29331829, 264908, 264909, 55182435, 265006, 264912, 264910, 60170631, 69043336, 265011, 265018, 18108351, 264448, 264288, 264766, 52844229, 21906765, 29148784, 65274791, 264456, 55182323, 60170394, 264558, 60432113, 264565, 264486, 264567, 264488, 56182575, 22278995, 33656286, 22278997, 22278998, 22278999, 264909, 264259, 29331822, 29331824, 66714117, 29331827, 35096052, 254107, 264905, 265003, 265009, 60432229, 33657402, 60434303, 55812038, 21906754, 65636542, 264910, 264911, 37168518, 265017, 265018, 265019, 264558, 264908, 264909, 265011, 21906767, 21906768, 55811957, 35096051, 265020, 60170615, 264909, 264691, 264652, 33657023, 264693, 65274620, 33657109, 18108370, 18108374, 265076, 35096423, 35096855, 264555, 264556, 18108381, 55182323, 60170394, 83373044, 18108385, 55526488, 80432113, 22279002, 264907, 265018, 264681, 264865, 264869
2936	87055663 (5871, 5872)	Novel Protein sim. GBank g1153863 (AC005569) - determined by GENSCAN prediction and mapped EST; match to EST R84329 (NC 942725) [Homo sapiens]	Contains protein domain (PF00836) - nuclease SET domain		
2937	84835096 (5873, 5874)	Novel Protein sim. GBank g15174409 (eNF_006101_1) (C122B - C122 antigen (Cytoplasmic tail-binding protein 2	UNCLASSIFIED		
2938	35419773 (5875, 5876)	Novel Protein sim. GBank g1331950 (lembCAA78720) - (Y11267) ubiquitin-conjugating enzyme [Mus musculus]	Contains protein domain (PF00179) - ubiquitin-conjugating enzyme		
2939	87786222 (5877, 5878)	Novel Protein sim. GBank g1397960 (lembCAA99629) - (275547) similar to WD domain, C-beta repeat, cDNA EST yk371b7.5 comes from this gene; cDNA EST yk31211.5 comes from this gene; cDNA EST yk46605.5 comes from this gene; cDNA EST yk47204.5 comes from this gene; cDNA EST yk25268.....	Contains protein domain (PF00400) - ATPase associated WD domain, C-beta repeat		

2946	84317315 (5897, 5892)	Novel Protein sim. GBank gl 5441953gb A043185.1 (AF072864) peroxisomal membrane protein PMP 24 [Homo sapiens]	UNCLASSIFIED	264488, 264239, 264508, 264509, 264906, 264907, 264909, 264510, 264511, 265007, 264512, 264910, 264591, 264593, 1810351, 264764, 264288, 264684, 264765, 285021, 264692, 33657109, 264622, 264629, 18103374, 264631, 264634, 264636, 264637, 18103390, 264638, 264639, 8337304, 224585, 264586, 264646, 264597, 22371989, 22778986, 22778987, 22778988, 22778989, 264648, 264599, 60432225, 21909764, 265008, 265009, 60432225, 21909764, 265010, 265017, 265019, 264766, 264685, 21906765, 21906766, 21906767, 21906768, 21906769, 265020, 264628, 18108370, 264629, 264630, 18108387, 80432113, 52646842, 22278995, 2644550, 26331854, 29331855, 29331827, 29331830, 264808, 265007, 265009, 265019, 264763, 264684, 264288, 264685, 264686, 21900767, 264691, 264692, 264693, 18108374, 55811576, 18108385, 22277602, 2644563, 2644567, 22278986, 22278987, 60432229, 29331826, 29331827, 29331828, 35696052, 29146489, 264104, 264107, 264993, 60712692, 264908, 264909, 264910, 264911, 264912, 264913, 52642220, 58181582, 21909767, 21909768, 21909769, 265022, 60170915, 33657023, 35696423, 263981, 264555, 80432113, 22277002
2947	87362952 (5895, 5894)	Novel Protein sim. GBank gl 3540281 gb A034383.1 - (AF056116) Aik-1 related protein [Fugu rubripes]	UNCLASSIFIED	264488, 264239, 264508, 264509, 264906, 264907, 264909, 264510, 264511, 265007, 264512, 264910, 264591, 264593, 1810351, 264764, 264288, 264684, 264765, 285021, 264692, 33657109, 264622, 264629, 18103374, 264631, 264634, 264636, 264637, 18103390, 264638, 264639, 8337304, 224585, 264586, 264646, 264597, 22371989, 22778986, 22778987, 22778988, 22778989, 264648, 264599, 60432225, 21909764, 265008, 265009, 60432225, 21909764, 265010, 265017, 265019, 264766, 264685, 21906765, 21906766, 21906767, 21906768, 21906769, 265020, 264628, 18108370, 264629, 264630, 18108387, 80432113, 52646842, 22278995, 2644550, 26331854, 29331855, 29331827, 29331830, 264808, 265007, 265009, 265019, 264763, 264684, 264288, 264685, 264686, 21900767, 264691, 264692, 264693, 18108374, 55811576, 18108385, 22277602, 2644563, 2644567, 22278986, 22278987, 60432229, 29331826, 29331827, 29331828, 35696052, 29146489, 264104, 264107, 264993, 60712692, 264908, 264909, 264910, 264911, 264912, 264913, 52642220, 58181582, 21909767, 21909768, 21909769, 265022, 60170915, 33657023, 35696423, 263981, 264555, 80432113, 22277002
2948	87626527 (5895, 5896)	Novel Protein sim. GBank gl 5666814 gb AB56554.2 - (AF001533) mitogen-induced [Mus musculus]	UNCLASSIFIED	264488, 264239, 264508, 264509, 264906, 264907, 264909, 264510, 264511, 265007, 264512, 264910, 264591, 264593, 1810351, 264764, 264288, 264684, 264765, 285021, 264692, 33657109, 264622, 264629, 18103374, 264631, 264634, 264636, 264637, 18103390, 264638, 264639, 8337304, 224585, 264586, 264646, 264597, 22371989, 22778986, 22778987, 22778988, 22778989, 264648, 264599, 60432225, 21909764, 265008, 265009, 60432225, 21909764, 265010, 265017, 265019, 264766, 264685, 21906765, 21906766, 21906767, 21906768, 21906769, 265020, 264628, 18108370, 264629, 264630, 18108387, 80432113, 52646842, 22278995, 2644550, 26331854, 29331855, 29331827, 29331830, 264808, 265007, 265009, 265019, 264763, 264684, 264288, 264685, 264686, 21900767, 264691, 264692, 264693, 18108374, 55811576, 18108385, 22277602, 2644563, 2644567, 22278986, 22278987, 60432229, 29331826, 29331827, 29331828, 35696052, 29146489, 264104, 264107, 264993, 60712692, 264908, 264909, 264910, 264911, 264912, 264913, 52642220, 58181582, 21909767, 21909768, 21909769, 265022, 60170915, 33657023, 35696423, 263981, 264555, 80432113, 22277002
2949	88175545 (5897, 5896)	Novel Protein sim. GBank gl 1332923 gb S87133 - probable membrane protein YOR204w - yeast (Saccharomyces cerevisiae)	UNCLASSIFIED	264488, 264239, 264508, 264509, 264906, 264907, 264909, 264510, 264511, 265007, 264512, 264910, 264591, 264593, 1810351, 264764, 264288, 264684, 264765, 285021, 264692, 33657109, 264622, 264629, 18103374, 264631, 264634, 264636, 264637, 18103390, 264638, 264639, 8337304, 224585, 264586, 264646, 264597, 22371989, 22778986, 22778987, 22778988, 22778989, 264648, 264599, 60432225, 21909764, 265008, 265009, 60432225, 21909764, 265010, 265017, 265019, 264766, 264685, 21906765, 21906766, 21906767, 21906768, 21906769, 265020, 264628, 18108370, 264629, 264630, 18108387, 80432113, 52646842, 22278995, 2644550, 26331854, 29331855, 29331827, 29331830, 264808, 265007, 265009, 265019, 264763, 264684, 264288, 264685, 264686, 21900767, 264691, 264692, 264693, 18108374, 55811576, 18108385, 22277602, 2644563, 2644567, 22278986, 22278987, 60432229, 29331826, 29331827, 29331828, 35696052, 29146489, 264104, 264107, 264993, 60712692, 264908, 264909, 264910, 264911, 264912, 264913, 52642220, 58181582, 21909767, 21909768, 21909769, 265022, 60170915, 33657023, 35696423, 263981, 264555, 80432113, 22277002
2850	85036870 (5898, 5900)	Novel Protein sim. GBank gl 466102 gb P34629 QJ6_CAEEL - PUTATIVE AMINOPEPTIDASE ZK353.6 IN CHROMOSOME III	UNCLASSIFIED	264488, 264239, 264508, 264509, 264906, 264907, 264909, 264510, 264511, 265007, 264512, 264910, 264591, 264593, 1810351, 264764, 264288, 264684, 264765, 285021, 264692, 33657109, 264622, 264629, 18103374, 264631, 264634, 264636, 264637, 18103390, 264638, 264639, 8337304, 224585, 264586, 264646, 264597, 22371989, 22778986, 22778987, 22778988, 22778989, 264648, 264599, 60432225, 21909764, 265008, 265009, 60432225, 21909764, 265010, 265017, 265019, 264766, 264685, 21906765, 21906766, 21906767, 21906768, 21906769, 265020, 264628, 18108370, 264629, 264630, 18108387, 80432113, 52646842, 22278995, 2644550, 26331854, 29331855, 29331827, 29331830, 264808, 265007, 265009, 265019, 264763, 264684, 264288, 264685, 264686, 21900767, 264691, 264692, 264693, 18108374, 55811576, 18108385, 22277602, 2644563, 2644567, 22278986, 22278987, 60432229, 29331826, 29331827, 29331828, 35696052, 29146489, 264104, 264107, 264993, 60712692, 264908, 264909, 264910, 264911, 264912, 264913, 52642220, 58181582, 21909767, 21909768, 21909769, 265022, 60170915, 33657023, 35696423, 263981, 264555, 80432113, 22277002
2851	85036870 (5898, 5900)	Novel Protein sim. GBank gl 466102 gb P34629 QJ6_CAEEL - PUTATIVE AMINOPEPTIDASE ZK353.6 IN CHROMOSOME III	UNCLASSIFIED	264488, 264239, 264508, 264509, 264906, 264907, 264909, 264510, 264511, 265007, 264512, 264910, 264591, 264593, 1810351, 264764, 264288, 264684, 264765, 285021, 264692, 33657109, 264622, 264629, 18103374, 264631, 264634, 264636, 264637, 18103390, 264638, 264639, 8337304, 224585, 264586, 264646, 264597, 22371989, 22778986, 22778987, 22778988, 22778989, 264648, 264599, 60432225, 21909764, 265008, 265009, 60432225, 21909764, 265010, 265017, 265019, 264766, 264685, 21906765, 21906766, 21906767, 21906768, 21906769, 265020, 264628, 18108370, 264629, 264630, 18108387, 80432113, 52646842, 22278995, 2644550, 26331854, 29331855, 29331827, 29331830, 264808, 265007, 265009, 265019, 264763, 264684, 264288, 264685, 264686, 21900767, 264691, 264692, 264693, 18108374, 55811576, 18108385, 22277602, 2644563, 2644567, 22278986, 22278987, 60432229, 29331826, 29331827, 29331828, 35696052, 29146489, 264104, 264107, 264993, 60712692, 264908, 264909, 264910, 264911, 264912, 264913, 52642220, 58181582, 21909767, 21909768, 21909769, 265022, 60170915, 33657023, 35696423, 263981, 264555, 80432113, 22277002
2951	87392357 (5901, 5902)	Novel Protein sim. GBank gl 4668992 pmo CAB41450.1 - (AJ238248) centaurin beta2 [Homo sapiens]	UNCLASSIFIED	264488, 264239, 264508, 264509, 264906, 264907, 264909, 264510, 264511, 265007, 264512, 264910, 264591, 264593, 1810351, 264764, 264288, 264684, 264765, 285021, 264692, 33657109, 264622, 264629, 18103374, 264631, 264634, 264636, 264637, 18103390, 264638, 264639, 8337304, 224585, 264586, 264646, 264597, 22371989, 22778986, 22778987, 22778988, 22778989, 264648, 264599, 60432225, 21909764, 265008, 265009, 60432225, 21909764, 265010, 265017, 265019, 264766, 264685, 21906765, 21906766, 21906767, 21906768, 21906769, 265020, 264628, 18108370, 264629, 264630, 18108387, 80432113, 52646842, 22278995, 2644550, 26331854, 29331855, 29331827, 29331830, 264808, 265007, 265009, 265019, 264763, 264684, 264288, 264685, 264686, 21900767, 264691, 264692, 264693, 18108374, 55811576, 18108385, 22277602, 2644563, 2644567, 22278986, 22278987, 60432229, 29331826, 29331827, 29331828, 35696052, 29146489, 264104, 264107, 264993, 60712692, 264908, 264909, 264910, 264911, 264912, 264913, 52642220, 58181582, 21909767, 21909768, 21909769, 265022, 60170915, 33657023, 35696423, 263981, 264555, 80432113, 22277002

2952	95322952 (5903, 5904)	Novel Protein sim. GBank gll55653jrm(CAB51405.1) - (ALD96881) hypothelial protein [Homo sapiens]	Contains protein domain (PF00650) - transcriplator CRAL/TRIO domain.	264687, 526455, 21506766, 21906769, 267896, 265020, 264690, 60432049, 264259, 264683, 29331822, 18108365, 29331825, 60432289, 33857109, 18108368, 29331827, 35696052, 27486382, 264508, 264905, 20281149, 264906, 264907, 29331830, 264908, 264909, 35695855, 264511, 265008, 265009, 264910, 264635, 264636, 60432229, 264638, 60433356, 264639, 264756, 87168518, 265017, 22270000, 22279002, 264760, 264583, 264482, 18108351, 264446, 264288
2953	86036575 (5905, 5906)	Novel Protein sim. GBank gll19832jpe(P105S1ERC, BASIT - PROBABLE PHOSPHO-SERINE AMINOTRANSFERASE (PSAT), ENDOMETRIAL PROGESTERONE-INDUCED PROTEIN) (EPIP)	Contains protein domain (PF00266) - UNCLASSIFIED Aminotransferases class-V	8103395, 35894075, 22278996, 28331822, 29331825, 60432289, 33857109, 18108365, 264408, 264906, 264907, 264908, 264909, 35657402, 264585, 264586, 264591, 264594, 33657402, 264585, 264586, 264758, 52646317, 21008754, 33657094, 52644236, 87168559, 264600, 264760, 264681, 18108351, 264764, 264589, 264288, 264687, 21906766, 21906766, 21906787, 21906788, 21906769, 35695917, 33657023, 18108384, 52644126, 33657109, 33657349, 18108374, 263976, 35695855, 264637, 264638, 87188518, 264482, 264583, 264585, 29331822
2954	86086226 (5907, 5908)	Novel Protein sim. GBank gll46825jpe(NP_005251.1)-pGOF9 - growth differentiation factor 9	Contains protein domain (PF001019) - Transforming growth factor beta like domain	264259, 29331822, 29331824, 29331825, 29331826, 35696052, 264906, 52644045, 264912, 60432229, 265018, 265019, 52644126, 33657109, 33657109, 21906788, 21906789, 265021, 60170615, 58810794, 264667
2955	87698426 (5909, 5910)	Novel Protein sim. GBank gll3454j73 (AF064203) - serine/threonine protein kinase [HCO] [Rattus norvegicus]	Contains protein domain (PF00069) - Eucaryotic protein kinase domain	22278996, 264259, 29331827, 264908, 21906788
2956	85789745 (5911, 5912)	Novel Protein sim. GBank gll46825jpe(AAD27830.1)-AF121857 sorting nexin 7 [Homo sapiens]	Contains protein domain (PF01078) - PX domain	22278996, 264259, 29331827, 264908, 21906788
2957	90933301 (5913, 5914)	Novel Protein sim. GBank gll45032jpe(NP_000069.1)-pCPT2 - carnitine palmitoyltransferase II precursor	cadherin	22278996, 264259, 29331824, 29331827, 265008, 264585, 264758, 265010, 265011, 264448, 264783, 264683, 264286, 264685, 18108357, 29148629, 204690, 18108362, 264693, 18108370, 60431528, 18108374, 264634, 18108381, 56182323, 18108382, 18108395, 18108386, 34528488, 87168518, 264482, 264487, 264585, 264586, 264681, 264580, 264620, 264631, 264587
2958	87440014 (5915, 5916)	Novel Protein sim. GBank gll42425jpe(BA374007.1) - (AB220891) KIA0984, protein [Homo sapiens]		

2559	95109420 (5917, 5918)	[Novel Protein sim. GBank g1958221 (U33005) - Tbc1 [Mus musculus]	Contains protein domain (PF00566) - oncogene TBC domain	263994, 22278997, 264259, 64432049, 29331826, 29331826, 33657402, 68712502, 56183436, 35690052, 265008, 264509, 60433354, 60433438, 264598, 265017, 265018, 264603, 264288, 264756, 264769, 21906766, 21906767, 21906769, 265020, 60170615, 264692, 27486285, 18108374, 80274791, 35698855, 83373044, 56526486, 60432113
2600	87420591 (5918, 5920)		UNCLASSIFIED	60170615, 35996423, 87168474, 265010, 87168518, 264483
2601	85131416 (5921, 5922)	[Novel Protein sim. GBank g15956560 (embJCAB05177.2) - (293258) protein using GeneIndex: similar to MYD domain, G-beta repeats [Caenorhabditis elegans]	Contains protein domain (PF00400) - MYD domain, G-beta repeat	22278997, 22278998, 264259, 29331822, 29331824, 29331826, 29331828, 264907, 264808, 52944045, 265008, 33657402, 21906794, 87168474, 265011, 87168599, 265017, 21906769, 265020, 60170615, 264692, 27486285, 18108374, 80274791, 35698855, 18108374, 56526486, 60432113, 18108385, 87168518, 22279002, 264554, 264587
2602	87512700 (5923, 5924)		UNCLASSIFIED	35698856, 22278997, 264092, 264094, 264259, 29331824, 60714117, 29331825, 60432289, 29331826, 29331827, 29331828, 35698852, 264508, 264905, 264509, 264907, 264908, 264909, 264510, 264512, 264593, 264594, 60433438, 284758, 52646317, 264602, 264603, 264605, 264760, 264762, 264764, 264288, 264766, 264686, 264768, 264769, 35698917, 265020, 264691, 264694, 264636, 264837, 264838, 264635, 18108385, 264583, 264585, 264586, 264597, 264486
2603	95313464 (5923, 5925)	[Novel Protein sim. GBank g1240230 (b19AX74590.1) - (AB020674) KIAA0857 protein [Homo sapiens]	Contains protein domain (PF00010) - Helix-loop-helix DNA-binding domain	264587, 264589, 264591, 22278998, 29331827, 29331828, 264509, 264512, 29331828, 265007, 265008, 264592, 264594, 21906754, 265018, 264769, 264687, 29148627, 29148784, 265020, 33657023, 264693, 65274620, 33657182, 27486281, 264693, 55910764, 35698423, 264555, 264636, 264637, 264557, 264558, 264563, 264259, 29331828, 33657402, 265017, 265018, 264692, 18108386, 35698423, 83373044, 18108388
2604	94324611 (5927, 5928)		UNCLASSIFIED	

2965	6034762 (5929, 5930)	Novel Protein sim. GBank g14665447(1e1p) - Keesle (mouse) mal- related leucine zipper homolog	Transcript factor	264235, 26331826, 264500, 264509, 264805, 264907, 264908, 264909, 26491, 265008, 264910, 264991, 264993, 264994, 264995, 265011, 264700, 264762, 264764, 264788, 264885, 264766, 264692, 264692, 264628, 264629, 264695, 264630, 264631, 264632, 264634, 264635, 264636, 264637, 264638, 264639, 264633, 264637, 264638, 6043289, 264682, 264448
2966	91725245 (5931, 5932)	Novel Protein sim. GBank g1526275 (1e1p) - Rabin3 (Xenopus laevis)		UNCLASSIFIED
2967	94658303 (5933, 5934)	Novel Protein sim. GBank g1624225 (U1915) - Rabin3 (Rattus norvegicus)		UNCLASSIFIED
2968	95302776 (5935, 5936)	Novel Protein sim. GBank g1429713(g1429713) - COH-123 protein (Homo sapiens)	Contains protein domain Zinc finger, C3HC4 type (RING finger)	264488, 264508, 264509, 264508, 264509, 264511, 264510, 264594, 264755, 2648542, 264762, 264764, 265021, 264556, 1810831, 264564, 264486 264687, 2645156, 21906765, 52645365, 21906767, 18108358, 36568423, 22278996, 36568286, 22278997, 265020, 22278999, 265021, 265022, 264093, 264636, 264690, 52644150, 264259, 33657023, 52646080, 264693, 29331822, 55182181, 29331824, 66714117, 29331825, 33109954, 52645129, 29331826, 21900754, 33657182, 29331827, 91768474, 265010, 87168595, 265018, 29331828, 36569052, 27482282, 87168518, 29331829, 55182183, 22279002, 264563, 18108361, 264696, 264696, 264696, 66712502, 264565, 264696, 264338, 52646842, 22278996, 22278998, 22278998, 60432049, 264259, 29331824, 29331825, 29331826, 29331828, 264509, 264699, 52644045, 56182435, 265009, 60434338, 55812038, 21900754, 265011, 87166559, 265018, 265019, 264448, 264488, 264369, 52644229, 21900766, 21906768, 21906769, 29148794, 265020, 265021, 52644150, 264691, 33557109, 18108374, 56182323, 60170394, 87168518, 60432113, 22279000
2969	95310657 (5937, 5938)	Novel Protein sim. GBank g13024743(g13024743) - THERMOSOME, ALPHA SUBUNIT (CHAPERONIN ALPHA SUBUNIT)	eph	UNCLASSIFIED
2970	88038071 (5939, 5940)	Novel Protein sim. GBank g13165407 (AC004765) - Homo sapiens	Contains protein domain Homeobox domain	UNCLASSIFIED

2974	91215301 (5943, 5944)	Novel Protein sim. GBank g1728337/p19184JALU7_HUMAN - IIII ALU SUBFAMILY SQ WARNING ENTRY IIII	in7	254488, 36182575, 35696286, 5699007, 5, 29331824, 29331826, 29146499, 264508, 264905, 264907, 264112, 264910, 21906794, 87168559, 265018, 265019, 18108351, 264689, 21908765, 21906767, 21006768, 265020, 265021, 80170615, 18108364, 264638, 264629, 26106374, 264636, 264556 264558, 83373044, 18108384, 18108385, 87168518, 264594, 264567 265017, 35693917, 265021, 33537109, 22276002, 264663
2972	86625943 (5943, 5944)	Novel Protein sim. GBank g1728337/p19184JALU7_HUMAN - IIII ALU SUBFAMILY SQ WARNING ENTRY IIII	kinase	254488, 36182575, 35696286, 5699007, 5, 29331824, 29331826, 29146499, 264508, 264905, 264907, 264112, 264910, 21906794, 87168559, 265018, 265019, 18108351, 264689, 21908765, 21906767, 21006768, 265020, 265021, 80170615, 18108364, 264638, 264629, 26106374, 264636, 264556 264558, 83373044, 18108384, 18108385, 87168518, 264594, 264567 265017, 35693917, 265021, 33537109, 22276002, 264663
2973	91215301 (5943, 5944)	Novel Protein sim. GBank g1728337/p19184JALU7_HUMAN - IIII ALU SUBFAMILY SQ WARNING ENTRY IIII	UNCLASSIFIED	254488, 36182575, 35696286, 5699007, 5, 29331824, 29331826, 29146499, 264508, 264905, 264907, 264112, 264910, 21906794, 87168559, 265018, 265019, 18108351, 264689, 21908765, 21906767, 21006768, 265020, 265021, 80170615, 18108364, 264638, 264629, 26106374, 264636, 264556 264558, 83373044, 18108384, 18108385, 87168518, 264594, 264567 265017, 35693917, 265021, 33537109, 22276002, 264663
2974	91273002 (5947, 5948)	Novel Protein sim. GBank g1766117 (L1834) - nuclear protein [Ensis minor]	UNCLASSIFIED	254488, 36182575, 35696286, 5699007, 5, 29331824, 29331826, 29146499, 264508, 264905, 264907, 264112, 264910, 21906794, 87168559, 265018, 265019, 18108351, 264689, 21908765, 21906767, 21006768, 265020, 265021, 80170615, 18108364, 264638, 264629, 26106374, 264636, 264556 264558, 83373044, 18108384, 18108385, 87168518, 264594, 264567 265017, 35693917, 265021, 33537109, 22276002, 264663
2975	85325213 (5948, 5950)	Novel Protein sim. GBank g13800812 (p19184JALU7_HUMAN - IIII ALU SUBFAMILY SQ WARNING ENTRY IIII) cDNA EST Y448010.5 comes from this gene [Ctenorhabdalis elegans]	Contains protein domain (PF00632) - ubiquitin - HECT domain (ubiquitin- transferase).	254488, 36182575, 35696286, 5699007, 5, 29331824, 29331826, 29146499, 264508, 264905, 264907, 264112, 264910, 21906794, 87168559, 265018, 265019, 18108351, 264689, 21908765, 21906767, 21006768, 265020, 265021, 80170615, 18108364, 264638, 264629, 26106374, 264636, 264556 264558, 83373044, 18108384, 18108385, 87168518, 264594, 264567 265017, 35693917, 265021, 33537109, 22276002, 264663
2976	87771202 (5951, 5952)	Novel Protein sim. GBank g1567913 (p19184JALU7_HUMAN - IIII ALU SUBFAMILY SQ WARNING ENTRY IIII) B-DNA L141488 [Drosophila melanogaster]	transport	254488, 36182575, 35696286, 5699007, 5, 29331824, 29331826, 29146499, 264508, 264905, 264907, 264112, 264910, 21906794, 87168559, 265018, 265019, 18108351, 264689, 21908765, 21906767, 21006768, 265020, 265021, 80170615, 18108364, 264638, 264629, 26106374, 264636, 264556 264558, 83373044, 18108384, 18108385, 87168518, 264594, 264567 265017, 35693917, 265021, 33537109, 22276002, 264663
2977	91725254 (5955, 5956)	Novel Protein sim. GBank g1567913 (p19184JALU7_HUMAN - IIII ALU SUBFAMILY SQ WARNING ENTRY IIII) B-DNA L141488 [Drosophila melanogaster]	UNCLASSIFIED	254488, 36182575, 35696286, 5699007, 5, 29331824, 29331826, 29146499, 264508, 264905, 264907, 264112, 264910, 21906794, 87168559, 265018, 265019, 18108351, 264689, 21908765, 21906767, 21006768, 265020, 265021, 80170615, 18108364, 264638, 264629, 26106374, 264636, 264556 264558, 83373044, 18108384, 18108385, 87168518, 264594, 264567 265017, 35693917, 265021, 33537109, 22276002, 264663

2978	87332059 (5956, 5956)	Novel Protein sim. GBank gll746548 (U23322) - No definition line found [Caenorhabditis elegans]	Contains protein domain (PF00480) - UNCLASSIFIED RCK family	22278995, 22278996, 22278997, 22278998, 22278999, 23031027, 25146499, 56102435, 265006, 265007, 265009, 60433396, 60433438, 21906754, 265010, 265011, 265017, 265018, 265019, 264488, 264685, 264688, 21906765, 21906768, 21906767, 21906769, 265020, 255021, 265022, 35996423, 264639, 60432113, 22279000, 22279002
2979	91725354 (5957, 5958)	Novel Protein sim. GBank gll56275 [emb]CAB-5680.11- (AJ431171) Xenopus RPA interacting protein alpha [Xenopus laevis]	complement	264488, 65274572, 35994075, 22278999, 255022, 35996423, 264639, 60432113, 22279000, 22279002, 55811927, 33055793, 264905, 60432113, 264482, 55811926, 55182323, 60432113, 264482, 265009, 21906767, 265081, 22279000
2980	86236600 (5959, 5960)			
2981	87376330 (5961, 5962)			
2982	95303673 (5963, 5964)	Novel Protein sim. GBank gll4928767 [h]AD34144.1 (AF151907) CGI-149 protein [Homo sapiens]	UNCLASSIFIED	22278995, 56994075, 22278996, 22278997, 22278998, 22278999, 264492, 20331834, 29331827, 29331828, 264905, 264591, 264592, 264594, 264595, 264596, 33657094, 264448, 21906765, 21906766, 21906767, 21906768, 21906769, 265020, 265022, 18108395, 33657182, 33657349, 35996423, 60432113, 22279000, 22279002
2983	91725354 (5965, 5966)	Novel Protein sim. GBank gll56275 [emb]CAB-5680.11- (AJ431171) Xenopus RPA interacting protein alpha [Xenopus laevis]		60424178, 52548842, 18108396, 22278997, 264905, 60432049, 264259, 29331822, 60432289, 33555910, 264905, 52644045, 265009, 60431735, 8716881, 265016, 265019, 10109180, 21906765, 21906768, 35996591, 33657023, 52544129, 18108370, 35996423, 83313044, 56529465, 60432113, 264404, 22278902
2984	94136467 (5967, 5968)	Novel Protein sim. GBank gll2393734 AC002542 - similar to C. elegans F1A10.5, 80% similarity to Z58297 (PID g1130618) [Homo sapiens]	ATPase associated	
2985	87099607 (5969, 5970)	Novel Protein sim. GBank gll103160 [p]IS21.26 - finger protein unkempt - fruit fly [Drosophila melanogaster]	UNCLASSIFIED	264810, 55812038, 36181562, 55811997, 264628, 55810794, 264632, 264635, 60432113
2986	86284651 (5971, 5972)			55811997, 264595
2987	86455934 (5973, 5974)		UNCLASSIFIED	22278997, 264595

2997	87627440 (5993, 5994) Novel Protein sim. GBank gi 545652:cdj BAA7648.1 - (AB023221) KIAA1004 protein [Homo sapiens]	homeobox	264488, 56182575, 264259, 6671417, 2931826, 35696052, 264508, 264509, 264497, 264900, 265006, 87163474, 265019, 264448, 264682, 264685, 264766, 21908764, 21908766, 21900768, 21908769, 27486261, 18103374, 35696423, 264634, 264635, 264636, 264557, 18108385, 2718518
2998	88050381 (5995, 5996) Novel Protein sim. GBank gi 3047459:emb CA02242 - (AL034364) cDNA EST x2595.3 comes from this gene; cDNA EST y62590.5 comes from this gene; cDNA EST EMBL_M75923 comes from this gene [Caenorhabditis elegans]	UNCLASSIFIED	32645395, 22278997, 264508, 264996, 18103374, 21900765, 21900767, 18103370, 18103374, 35696423, 264636, 264639
2999	94847055 (5997, 5998) Novel Protein sim. GBank gi 1540819:pf P18835 CC19, CAEL - CUTICLE COLLAGEN 19	UNCLASSIFIED	56182575, 22278996, 29147620, 26331825, 29146498, 29146499, 264905, 66712502, 265006, 265009, 21906754, 85655542, 18103351, 29148627, 29148629, 60170615, 33657109, 27486262, 18103370, 18103374, 264556, 264557, 264558, 60170394, 18103385, 264553
3000	55595370 (5999, 6000) Novel Protein sim. GBank gi 1163174 (U3257) - similar to Yeast Scdip, Swiss-Prot Accession Number P32844; similar to mammalian tRNA ^{Ala} , Swiss-Prot Accession Number Q01169; Method: conceptual translation supplied by author [Rattus norvegicus]	UNCLASSIFIED	264487, 22278997, 22278999, 264259, 28331822, 29331824, 35696052, 29146498, 264508, 264905, 264906, 264907, 264908, 264909, 264510, 264511, 265006, 265007, 265008, 265009, 264910, 33657402, 264757, 264535, 264599, 264758, 21906754, 265011, 264600, 265017, 265018, 264605, 265019, 264760, 264761, 264762, 264681, 264882, 264883, 264884, 264885, 264886, 264887, 264768, 264769, 21906765, 21906768, 35659917, 265020, 264591, 264692, 33657023, 264693, 33657109, 33657182, 27488281, 264828, 264829, 18103374, 18103376, 35696423, 35659555, 264630, 264631, 264632, 264634, 264635, 264636, 264637, 264556, 264638, 264639, 60170394, 83373044, 20798451, 22279002, 264583, 264486, 264587
3001	88778445 (6001, 6002) Novel Protein sim. GBank gi 2075470 (AC020273) - Purine gene, human predictions confirmed by EST analysis [Homo sapiens] (X12205) (ND.gi1679648), D31582 (ND.gi2644445), (ND.gi1733515), R55640 (ND.gi3003335) and F10621 (ND.gi709111) [Homo sapiens]	cathepan	18103394, 52646844, 56182575, 29331824, 29331825, 29331827, 264910, 33109994, 5264298, 265017, 265018, 264288, 265020, 265021, 52644109, 264694, 35695763, 53810794, 35696423, 56182323, 18103387, 264488, 29331824, 33331825, 29331826, 29331827, 29331828, 264905, 264910, 265009, 21906754, 264682, 264688, 33657023, 264565
3002	87718167 (6003, 6004) Novel Protein sim. GBank gi 3599478 (AF085189) - Nycteh [Acanthamoeba castellanii]	UNCLASSIFIED	

3013	91238799 (6025, 6026)	Novel Protein sim. gBank gi18702236 (AC003787) - R33374_1 [Homo sapiens]	Contains protein domain (PF00000) - Transcription factor domain, G-beta repeat	264468, 263994, 35698286, 22278997, 264259, 25331824, 60424665, 66714117, 35595052, 264905, 264906, 264907, 264908, 264909, 56182435, 264511, 264512, 264910, 264991, 264592, 264593, 264594, 33657402, 60433438, 284595, 264596, 5812038, 264758, 33109954, 21906754, 265010, 265018, 264604, 264760, 264682, 264683, 264764, 264369, 264288, 264765, 264766, 264696, 264768, 264687, 21906757, 35695917, 265020, 33657023, 264682, 264693, 33657109, 264628, 264629, 35811976, 35696425, 35695865, 264630, 264631, 35695866, 264632, 264633, 264634, 35695867, 264635, 264636, 264637, 264638, 264639, 35373044, 264683, 264595, 264586, 264587
3014	79877263 (6027, 6028)	Novel Protein sim. gBank gi1878374 (mbjCA03081) - (Z68879) Similarly to Yeast Cht12p protein (PIR Acc. No. S64453); cDNA EST EMBL D27850 comes from this gene; cDNA EST EMBL D27849 comes from this gene; cDNA EST EMBL D33447 comes from this gene; cDNA EST EMBL D33316 comes from ..	ATPase associated	264750
3015	56956461 (6029, 6030)			22278995, 22278996, 22278997, 264259, 25331824, 29331828, 264906, 265007, 265008, 264910, 265011, 265017, 265019, 264991, 33657109, 18108370, 35695855, 264596, 264594
3016	87759945 (6031, 6032)	Novel Protein sim. gBank gi1168819 (p41733) COPI _ YEAST - CELL DIVISION CONTROL PROTEIN 91	UNCLASSIFIED	52644807, 52644808, 56954075, 52644580, 25331822, 29331824, 35696062, 33659970, 264905, 264906, 33657084, 265017, 265019, 52644232, 33657109, 27468261, 52644159, 33657023, 33657109, 27468261, 27468262, 27468264, 33657340, 27468265, 35698763, 35698655, 87188518
3017	95011154 (6033, 6034)	Novel Protein sim. gBank gi1558965 (bj03AAV6851.1) - (AB023224) KIAA1007 protein [Homo sapiens]		264485, 18103397, 22278996, 35696266, 22278999, 264259, 29331822, 60432299, 264908, 25331830, 264905, 56182435, 265006, 265007, 265008, 265009, 264991, 60433358, 60433438, 52646317, 21906754, 55811386, 265010, 265011, 87168559, 265017, 265018, 265019, 264288, 264687, 21906765, 21906766, 21906767, 21906769, 265020, 265022, 65274620, 52645129, 33657109, 33657162, 18108370, 263972, 18108374, 264631, 52644332, 53373044, 18108381, 18108384, 35658495, 87168518, 264404, 6043113, 22278995, 264587

3018	11073891 (6035, 6036)	Novel Protein sim. GBank g13215332 (A0040200) - Unknown gene product (Homo sapiens)	oncogene	264558 264565, 5864507, 18106394, 63274572, 56182575, 22278994, 22278995, 56994075, 22278996, 22278999, 364259, 29331822, 29331824, 60432289, 39331827, 264908, 56182435, 265007, 265009, 60432229, 264593, 60433356, 55812036, 21060754, 87184474, 265011, 87166559, 265017, 265018, 265019, 264681, 18108351, 264448, 264682, 264683, 11035354, 264685, 264687, 264689, 21906766, 21906768, 21906769, 32844150, 264690, 264691, 33657023, 264692, 264693, 33657019, 264694, 33657020, 264695, 65571365, 264696, 58644332, 58183233, 18108365, 87166518, 22279000, 22279002, 264563
3020	9431251 (6039, 6040)	Novel Protein sim. GBank g13414839 (AF061529) - [Mus musculus] Regulator of chromosome condensation (RCC1)	Contains protein domain (PF00415) - ATPase associated Regulator of chromosome condensation (RCC1)	264483, 263594, 3599286, 364259, 264508, 264905, 264909, 264906, 264907, 264908, 264909, 264910, 264910, 60174639, 264900, 264603, 264760, 264752, 264682, 264763, 264764, 264288, 264359, 264766, 264687, 264698, 264769, 55811957, 35695917, 33657023, 264628, 35996423, 35695955, 264630, 264632, 264634, 264635, 264636, 264637, 264636, 264637, 264638, 264639, 63373044, 18108385, 264564, 264567, 264485 264769, 264629, 264482
3021	6047512 (6041, 6042)	Novel Protein sim. GBank g1360889 (m1CAB93006) - (Z95559) cDNA EST Y42364.5 comes from this gene. cDNA EST EMBL C13455 comes from this gene. cDNA EST Y4329g6.5 comes from this gene. cDNA EST CEMSH45R comes from this gene [Caenorhabditis elegans]	UNCLASSIFIED	264259, 29331826, 29331828, 264288, 264566
3022	87716500 (6043, 6044)	Novel Protein sim. GBank g1416992 (p32323) YG41, YEAST - A-AGGLUTININ ATTACHMENT SUBUNIT PRECURSOR	Contains protein domain (PF00614) - Phosphodiesterase D, Active site motif	UNCLASSIFIED
3023	9530484 (6045, 6046)	Novel Protein sim. GBank g1395671 (L11275) - selected as a weak suppressor of the growth defect of the dependent RNA polymerase I and II [Saccharomyces cerevisiae]	UNCLASSIFIED	60432049, 264680, 21906760, 53811957, 35995917, 264680, 264555, 264559 264593, 55815716
3024	86673305 (6047, 6048)			
3025	65705629 (6048, 6050)			

3026	87643662 (6051, 6052)	Novel Protein sim. GBank gi3024052[gb]/7824(KARI_RAT) - KALIRIN (PAM COOH-TERMINAL INTERACTOR PROTEIN 10) (P-CIP10)		UNCLASSIFIED	22276986, 22276997, 264490, 26331825, 264111, 265007, 60170831, 265010, 87168359, 265019, 21066765, 29146827, 263997, 20281149, 20281068, 263975, 263977, 20281071, 56526486, 22279000
3027	84844563 (6053, 6054)	Novel Protein sim. GBank gi4928547[gb]/A0230484.1(AF151847) COL1A9 protein [Homo sapiens]	Contains protein domain (PF01529) - DHC zinc finger domain	UNCLASSIFIED	22276987, 22276996, 22276996, 35696286, 18106334, 22276996, 60432049, 264259, 29331822, 29331824, 66714117, 26331825, 60432399, 29331826, 29331827, 29331828, 35696052, 264905, 264907, 29331830, 264908, 264909, 264510, 265007, 265008, 265009, 264910, 33657402, 264566, 21905754, 265010, 265011, 37168359, 264907, 265010, 18106351, 264908, 264909, 264910, 264911, 264912, 264913, 264914, 264915, 264916, 264917, 264918, 264919, 264920, 21906765, 21906766, 21906767, 21906768, 21906769, 20148929, 35695917, 21906768, 21906769, 21906770, 264692, 33657023, 264693, 52645128, 33657108, 27488281, 18108374, 55811578, 35698423, 55274791, 264636, 264656, 264657, 264638, 60170394, 264639, 264658, 83373044, 18108385, 55526486, 22276900, 22279002, 22276985, 22276986, 22276987, 22276989, 264259, 29331824, 29331827, 35696052, 264908, 265007, 265008, 265009, 60170331, 21905754, 265011, 87168359, 265018, 264782, 264983, 264785, 264689, 21906765, 21905768, 21906769, 23140628, 35695917, 264908, 264909, 264910, 264911, 264912, 264913, 264914, 18108374, 55811578, 35698423, 55274791, 264636, 264656, 264657, 264638, 60170394, 22276900, 22279002, 264482, 264564
3028	94231997 (6055, 6056)	Novel Protein sim. GBank gi3080521[emb]/CAA186511 - (A1022599) hypothetical protein [Schistosoma haematophyllum]	Contains protein domain (PF00400) - WD domain, G-beta repeat	UNCLASSIFIED	22276987, 22276989, 29331827, 264905, 264509, 264909, 264510, 264511, 264512, 87168474, 265019, 18106351, 21906768, 264534, 264690, 264693, 263969, 18108370, 264558, 22279000, 22279002, 264482
3029	87618284 (6057, 6058)			UNCLASSIFIED	
3030	87544928 (6059, 6060)	Novel Protein sim. GBank gi3757726[emb]/CAA187821 - (A1022727) dJ8019.1 (oligodactyl receptor-like protein [HsBM1-1]) [Homo sapiens]	Contains protein domain (PF00001) - 7 transmembrane receptor (rhodopsin family)	-tm7	
3031	91677953 (6061, 6062)	Novel Protein sim. GBank gi4530587[gb]/A022105.1 - (AF132000) TADA1 protein [Homo sapiens]		UNCLASSIFIED	22276985, 22276987, 22276988, 22276989, 264259, 29331822, 29331824, 29331826, 29331827, 29331828, 35696052, 264908, 265018, 21906765, 21906766, 21906767, 21906768, 265021, 263974, 18108374, 264908, 265022, 22279000, 22279002, 22279003, 22279004, 22279005

3053	195550373 (6105, 6106)	Newel Protein sim. GBank gi 3947613 emc CAA19465.1 - (AL023828) cDNA EST EMBL M89008 comes from this gene. cDNA EST Y4282d3.5 comes from this gene [Caenorhabditis elegans]	UNCLASSIFIED	65374572, 76181666, 2278995, 35696286, 22278986, 264259, 60432289, 265008, 265009, 60433435, 21906754, 265010, 87168539, 264900, 265016, 265019, 264763, 264764, 264288, 21906765, 21906768, 21906769, 21906769, 35695917, 18100374, 26454743, 264630, 56102323, 22279000, 26454743
3054	86943510 (6107, 6108)	Newel Protein sim. GBank gi 107621 Ugi S50755 - hypothetical protein YSP-3 - Chlamydomonas reinhardtii	UNCLASSIFIED	35696286, 35696052, 20331810, 264908, 264909, 264512, 264910, 265017, 264904, 264766, 265020, 33657109, 264628, 35695955, 264638, 264564, 264566, 264486, 60424179, 65274517, 56182575, 35696286, 22278986, 22278989, 60432049, 264259, 60424289, 60432289, 35696052, 56182435, 265008, 265009, 60170831, 60432229, 60431735, 60433356, 264594, 60433438, 21906754, 55811366, 265011, 87168559, 265019, 18100351, 264683, 264288, 264389, 264689, 21906768, 55811957, 35696917, 60170615, 33657023, 6274600, 33657109, 35695763, 60431522, 18103374, 55811957, 35696917, 60431522, 18103374, 55811957, 35696917, 60431850, 18108381, 56182323, 60170394, 18108395, 60432113, 264564, 264565, 264566
3055	95550337 (6109, 6110)	Newel Protein sim. GBank gi 480655 gb AA027717.1 JAF13294 - (AF132942) CGI-08 protein (Homo sapiens)	transport	264488, 264569, 18106394, 52846842, 22278987, 22278998, 22278999, 264259, 66714117, 29331826, 29331827, 35696052, 264508, 264509, 264905, 264906, 264907, 264908, 264909, 265006, 264512, 265007, 265008, 265009, 264910, 33657402, 55812038, 264596, 264758, 265010, 265011, 265017, 265019, 264760, 18108351, 264762, 264763, 264764, 264288, 264766, 264689, 18108357, 264768, 264769, 264689, 21906765, 21906766, 21906767, 21906768, 21906769, 35695917, 265020, 265021, 264691, 264693, 33657109, 18108370, 264692, 264694, 264695, 264696, 35695943, 35695945, 264700, 264631, 264632, 264634, 264635, 264636, 264637, 264638, 18108351, 83373544, 18108355, 264639, 22279002, 264663, 264584, 264585, 264586, 264486, 264567
3056	91661639 (6111, 6112)	Newel Protein sim. GBank gi 728837 sp P29194 ALU7_HUMAN - III ALU SUBFAMILY SQ WARNING ENTRY III	glycoprotein	264488, 264569, 18106394, 52846842, 22278987, 22278998, 22278999, 264259, 66714117, 29331826, 29331827, 35696052, 264508, 264509, 264905, 264906, 264907, 264908, 264909, 265006, 264512, 265007, 265008, 265009, 264910, 33657402, 55812038, 264596, 264758, 265010, 265011, 265017, 265019, 264760, 18108351, 264762, 264763, 264764, 264288, 264766, 264689, 18108357, 264768, 264769, 264689, 21906765, 21906766, 21906767, 21906768, 21906769, 35695917, 265020, 265021, 264691, 264693, 33657109, 18108370, 264692, 264694, 264695, 264696, 35695943, 35695945, 264700, 264631, 264632, 264634, 264635, 264636, 264637, 264638, 18108351, 83373544, 18108355, 264639, 22279002, 264663, 264584, 264585, 264586, 264486, 264567

3057	95412746 (6113, 6114)	Novel Protein sim. GBank glij39751 (9mb/C4.88660) - (249068) similar to GTP-binding protein: cDNA EST EMBL M83111 comes from this gene: cDNA EST EMBL D27703 comes from this gene: cDNA EST EMBL D27708 comes from this gene: cDNA EST EMBL D73788 comes from this gene: cDNA EST y4353.		struct	264508, 264505, 264507, 264508, 264509, 264510, 264512, 264510, 264502, 264594, 264767, 18108374, 264633, 264555, 264637, 264639, 264503, 264554, 264565, 264486
3058	78646226 (6115, 6116)			UNCLASSIFIED	264483
3059	87529425 (6117, 6118)	Novel Protein sim. GBank glij448803 (494AD5582.1A)F06287 - (AF092878) zinc RING finger protein SAG [Homo sapiens]	Contains protein domain (PF000097) - Zinc finger, C3HC4 type (RING finger)	UNCLASSIFIED	22278995, 22278996, 22278997, 22278998, 22278999, 264490, 264559, 29331824, 29331825, 29331827, 35696092, 29331828, 265007, 60433548, 265017, 265018, 265019, 264681, 264448, 264628, 264706, 21906765, 21906766, 21906767, 21906768, 29140629, 264636, 18108375, 264637, 264639, 18109385, 264563, 264567
3060	79346691 (6119, 6120)			UNCLASSIFIED	264567
3061	87740584 (6121, 6122)			UNCLASSIFIED	264112, 52644296, 21906768, 33657023, 263974, 18108395
3062	87619465 (6123, 6124)	Novel Protein sim. GBank glij445680 (94AD20663) - (AF070657) glutathione S-transferase subunit 13 homolog [Homo sapiens]		transfase	264500, 265008, 18108351, 264560
3063	80078023 (6125, 6126)	Novel Protein sim. GBank glij224532 (U53872) - ORF 73, contains large complex repeat CR 73 [Kaposi's sarcoma-associated herpesvirus]		UNCLASSIFIED	18108359, 264558
3064	91241526 (6127, 6128)	Novel Protein sim. GBank glij242031 (5b)BAAT7936.11 - (AB020720) KIAA0913 protein [Homo sapiens]	Contains protein domain (PF004003) - heavy-metal-associated domain	UNCLASSIFIED	52646385, 52646842, 6524572, 56182575, 55161686, 22278995, 22278996, 22278997, 22278999, 264259, 60432049, 29331824, 66714117, 264508, 264507, 264508, 56182435, 265009, 60432228, 60433548, 25912093, 264684, 264685, 264686, 264687, 264688, 264689, 21906768, 264691, 264692, 264693, 18109370, 18108377, 55811576, 264636, 56182323, 264558, 264639, 18108385, 22279000, 22279002
3065	91632021 (6129, 6130)	Novel Protein sim. GBank glij565674 (94AD45860.1A)C05066 - (AC005067) Supported by Human EST H98032.1 (MD.g87285A), mouse EST AA87042.1 (MD.g2965487), and ginsan [Homo sapiens]		UNCLASSIFIED	22278996, 22278998, 264593, 264594, 264595, 29331824, 60424265, 66714117, 264100, 264907, 265007, 264591, 60432229, 264593, 265011, 265019, 18108351, 264786, 264787, 21906765, 21906766, 264693, 20281069, 22279000, 22279002, 264492, 264566, 264567

3003	6809575 (6165, 6166)	Novel Protein sim. GBank gll68241 (U29468) - C56C10.3 gene product [Caenorhabditis elegans]	UNCLASSIFIED	264488, 264259, 29331823, 264105, 265008, 264591, 264592, 21906754, 264988, 264767, 21906768, 21906769, 29148784, 264691, 264632, 22279000
3004	87448568 (6167, 6168)	Novel Protein sim. GBank gll7677 (pJKA37475 - probable structural component p38 - borna disease virus		22278595, 60432289, 35696052, 2649035, 264506, 264907, 264508, 264908, 264909, 265006, 265007, 264910, 264593, 264595, 264758, 264369, 264288, 264788, 35095917, 265020, 18108374, 35096423, 264031, 264556, 264555, 264506, 264567, 264486
3005	87795781 (6169, 6170)	Novel Protein sim. GBank gll356597 (U60741) - CAG444 [Homo sapiens]	UNCLASSIFIED	265011, 264681
3006	87769942 (6171, 6172)	Novel Protein sim. GBank gll3894159 (AC055862) - hypothetical protein [Arabidopsis thaliana]	UNCLASSIFIED	22278989, 264092, 264255, 29331822, 29331825, 264108, 264112, 18108351, 18108352, 264591, 263974, 55810784, 233081, 264441
3007	87462598 (6173, 6174)			55810835, 56994075, 22278997, 22279008, 29331824, 29331825, 35696052, 60434348, 33109954, 21906754, 52646317, 285017, 264682, 264369, 264684, 21905767, 21906788, 265020, 264691, 33657023, 33657109, 52645129, 33657182, 27486202, 35695955, 87168518
3008	91224441 (6175, 6176)	Novel Protein sim. GBank gll353504 (AF001549) - Unknown gene product [Homo sapiens]	UNCLASSIFIED	264491
3009	95361242 (6177, 6178)	Novel Protein sim. GBank gll46861469 (AJ27182) - (AF077049) lamda-crystallin [Homo sapiens]	Contains protein domain (PF00725) - dehydrogenase 3-hydroxyacyl-CoA dehydrogenase	18108397, 65274572, 35192575, 5671686, 59994075, 35996286, 22278997, 22278998, 264259, 29331824, 29331825, 29331826, 29331828, 264907, 29331830, 264909, 56162435, 264510, 265007, 60170831, 98422229, 21906754, 55810841, 55810842, 264506, 264507, 264508, 264509, 264510, 264511, 264512, 264513, 264514, 264515, 264516, 56181562, 21905765, 21906786, 21906787, 21906788, 265021, 60170515, 27486282, 18108370, 60431528, 35696423, 264558, 264559, 60432113, 264468

[illegible]

3092	95314592 (6183, 6184)	Novel Protein sim. GBank glt1710260gP1580gP25; HUMAN - 40S RIBOSOMAL PROTEIN S2 (S4) (LUREP3 PROTEIN)	Cottaris protein domain (PF03333) Ribosomal protein S6	ribosomalprot	264848, 50424175, 18108396, 22278995, 58940755, 22278996, 35598266, 22278997, 22317898, 60412043, 264259, 29331822, 29331824, 29331825, 29331826, 29331827, 35896052, 29331828, 29146488, 29146489, 264508, 264509, 264905, 264906, 264907, 29331830, 264908, 264909, 264113, 264510, 264511, 265006, 264512, 265007, 265008, 264910, 265009, 60170531, 264591, 264592, 60431735, 264553, 264394, 60433438, 264595, 264758, 21900754, 265010, 265011, 264901, 264602, 265017, 264903, 264904, 265018, 264603, 265015, 264760, 264762, 264981, 18108351, 264763, 264682, 264448, 264764, 264683, 264765, 264369, 264766, 264768, 264686, 264767, 264687, 264768, 264769, 264688, 264770, 264689, 264771, 21900759, 20148629, 20148784, 35695917, 265020, 265021, 264534, 60170615, 264690, 264691, 264692, 65274520, 33657193, 27486262, 264628, 264629, 18108374, 263978, 18108377, 35686423, 264630, 264631, 264632, 264634, 264635, 264555, 264636, 264637, 264556, 264638, 264587, 264588, 264639, 60170594, 18105395, 264599, 60432229, 60431735, 264694, 264369, 264628, 264766, 21900707, 35598423, 63373044, 18108351, 264595, 18108356, 264259, 29331827, 29331828, 264398, 264767, 265002, 264691, 264693, 65274791, 56183323, 264584, 264585
3093	94318457 (6185, 6186)	Novel Protein sim. GBank glt0022587[emb]C:BA4347 11 - (Y17454) LSRF1 protein [Homo sapiens]	UNCLASSIFIED		
3094	94316675 (6187, 6188)	Novel Protein sim. GBank glt007346gP31c4[PB]_RAT - PHOSPHATIDYLETHANOLAMINE-BINDING PROTEIN (23 KD MORPHINE-BINDING PROTEIN) (P23K)	Contains protein domain (PF01161) - collagen (23 Phosphatidylethanolamine binding protein		

3095	94448162 (6189, 6190)	Novel Protein sim. GBank gH407759gBAAD31421 (JAF12444 - (AF12444)) MAGE tumor antigen D1 [Homo sapiens]	Contains protein domain (PF01454) - MAGE family	UNCLASSIFIED	18108397, 56182575, 22278995, 35696286, 56994075, 22278997, 22278999, 264239, 60432049, 60714117, 29331825, 60432289, 35696062, 33656970, 29146490, 264508, 264905, 264509, 29331830, 264509, 264510, 264511, 284512, 265007, 265008, 285009, 60170631, 264758, 21906754, 85658542, 265010, 265011, 87168556, 265017, 265018, 265019, 264760, 264681, 264682, 264683, 264764, 264389, 264288, 264688, 264768, 264769, 264689, 21906763, 21906766, 21906767, 56811687, 35696062, 264509, 264510, 264511, 264512, 264513, 264682, 33657023, 264693, 963972, 18108376, 56811576, 35696043, 264492, 60170394, 264639, 83373044, 18108385, 18108387, 65274727, 87168518, 60432113, 264482, 264583, 264584, 264586, 264487, 18108391, 22278995, 22278996, 22278997, 22278999, 29331824, 29331825, 29331826, 29331827, 33656970, 264905, 264906, 265008, 264910, 33657402, 265011, 265017, 265018, 264389, 21906766, 21906767, 21906768, 35696043, 265020, 60170815, 264691, 264692, 264693, 27486261, 27486262, 18108370, 80431528, 264634, 264635, 264639, 22279002, 264596, 264488, 29331822, 29331825, 60432289, 29331826, 264637, 264639, 264640, 264641, 264642, 55812033, 264643, 264644, 264645, 87158474, 87168559, 52844220, 21906765, 21906767, 18108376, 35696043, 52844332, 264638, 60432113, 22279002
3096	87756128 (6197, 6198)	Novel Protein sim. GBank gH8832221[db][BAA34470.1] - (AB018293) KIAA0750 protein [Homo sapiens]	Contains protein domain (PF00307) - Calponin homology (CH) domain	UNCLASSIFIED	264634, 264635, 264639, 22279002, 264596, 264488, 29331822, 29331825, 60432289, 29331826, 264637, 264639, 264640, 264641, 264642, 55812033, 264643, 264644, 264645, 87158474, 87168559, 52844220, 21906765, 21906767, 18108376, 35696043, 52844332, 264638, 60432113, 22279002
3097	88264865 (6193, 6194)	Novel Protein sim. GBank gH46828[db][CA037981] - (AL022508, 6272912, 1) PUTATIVE protein based on EST (matches) [Homo sapiens]	Contains protein domain (PF00646) - F-box domain.	UNCLASSIFIED	264634, 264635, 264639, 22279002, 264596, 264488, 29331822, 29331825, 60432289, 29331826, 264637, 264639, 264640, 264641, 264642, 55812033, 264643, 264644, 264645, 87158474, 87168559, 52844220, 21906765, 21906767, 18108376, 35696043, 52844332, 264638, 60432113, 22279002
3098	80258024 (6195, 6196)	Novel Protein sim. GBank gH03603[db][AA02145.1] - (D12821) cytochrome P-4501TBV [Homo sapiens]	cyto450	UNCLASSIFIED	264634, 264635, 264639, 22279002, 264596, 264488, 29331822, 29331825, 60432289, 29331826, 264637, 264639, 264640, 264641, 264642, 55812033, 264643, 264644, 264645, 87158474, 87168559, 52844220, 21906765, 21906767, 18108376, 35696043, 52844332, 264638, 60432113, 22279002
3100	87024241 (6199, 6200)	Novel Protein sim. GBank gH1083764[db][B46013 - prolins- rich proteoglycan 2 precursor, paroid - rat	UNCLASSIFIED	UNCLASSIFIED	264634, 264635, 264639, 22279002, 264596, 264488, 29331822, 29331825, 60432289, 29331826, 264637, 264639, 264640, 264641, 264642, 55812033, 264643, 264644, 264645, 87158474, 87168559, 52844220, 21906765, 21906767, 18108376, 35696043, 52844332, 264638, 60432113, 22279002
3101	79602134 (6201, 6202)		UNCLASSIFIED	UNCLASSIFIED	264634, 264635, 264639, 22279002, 264596, 264488, 29331822, 29331825, 60432289, 29331826, 264637, 264639, 264640, 264641, 264642, 55812033, 264643, 264644, 264645, 87158474, 87168559, 52844220, 21906765, 21906767, 18108376, 35696043, 52844332, 264638, 60432113, 22279002

3102	9122082 (6/03, 6/04)	Novel Protein sim. GBank gi3505709gb AA04781.1 (AF12835) Cytoplasmic phosphorolipid PACSIN2 (Homo sapiens)	Contains protein domain SH3 domain	struct	35696286, 22278986, 22278989, 28331827, 35696052, 264908, 264512, 265003, 28331833, 26433356, 33519954, 18108351, 264908, 264512, 265003, 264682, 264682, 33657023, 264638, 22279000, 264482, 264564 35695917, 264565
3103	90938004 (6/05, 6/06)	Novel Protein sim. GBank gi464564 jplP35392.1 (MUSE - RAS-RELATED PROTEIN RAB-17)		UNCLASSIFIED	
3104	87340633 (6/07, 6/08)	Novel Protein sim. GBank gi5032207 trf NP_005696	lptTSSC - tumor-suppressing STF cDNA 6	UNCLASSIFIED	264259, 264684, 264632, 33657182, 264558
3105	94148603 (6/08, 6/10)				22278997, 264259, 28331824, 35696052, 28331828, 264500, 264599, 264905, 264906, 264907, 264908, 264511, 264910, 264591, 264594, 264758, 264760, 264681, 264762, 264764, 264288, 264766, 264766, 264907, 33657023, 21906765, 21906769, 35695917, 33657023, 264638, 264638, 264638, 35695955, 264630, 264631, 264632, 264633, 264635, 264637, 264638, 264639, 33737044, 264404, 22278902, 264563, 264565, 264566, 264456, 264567
3106	95361416 (6/11, 6/12)	Novel Protein sim. GBank gi11938574 (U97180) - B0025.2 gene product [Caenorhabditis elegans]			22278996, 22278997, 22278998, 22278999, 264092, 264093, 264094, 28331822, 264906, 264907, 264908, 32844045, 56182435, 264112, 265008, 265009, 55812038, 265017, 265018, 264683, 264686, 264687, 264783, 52644229, 21906765, 21906768, 21906769, 55811957, 265020, 265022, 264690, 52644150, 264692, 264693, 18108370, 18108377, 55811576, 56182323, 18108385, 18108386, 22278900, 264583
3107	95343272 (6/13, 6/14)	Novel Protein sim. GBank gi3341441 emb C047635.1 (Y17794) winged-helix transcription factor [Gallus gallus]			22278985, 22278986, 35696286, 22278997, 264908, 264512, 265003, 264425, 28331832, 26433356, 26511829, 322289, 28331827, 28331828, 33658970, 264105, 264512, 265009, 60433356, 60433438, 265011, 265017, 265018, 21906765, 21906766, 21906767, 21906769, 265021, 264691, 33657108, 27488261, 27488265, 18108370, 263972, 18108374, 55811576, 18108385, 56556486, 264432, 264487, 56182435, 264288, 264680, 264564
3108	87340635 (6/15, 6/16)	Novel Protein sim. GBank gi5032207 trf NP_005696	lptTSSC - tumor-suppressing STF cDNA 6	UNCLASSIFIED	

[illegible]

3118	94565848 (6235, 6236)	Novel Protein sim. GBank g01380565jrmjCAB01444.11 - (Z70016) predicted using GeneFinder: similar to EGF-like domain EGF-like domain this gene [Caenorhabditis elegans]	Contains protein domain (PF00008) - Igf	52845156, 52846842, 56274572, 56182575, 22278995, 56994075, 22278996, 35969286, 22278997, 22278998, 22278999, 264229, 29331822, 29331824, 66741117, 29331826, 29331827, 35969052, 29331828, 264905, 264908, 29331830, 52844045, 56182435, 264510, 264511, 265007, 385000, 265009, 528439, 52846317, 21906754, 33657004, 265018, 264905, 265019, 264785, 264786, 264902, 264684, 264268, 264786, 56181182, 21906765, 21906766, 21906768, 21906769, 255020, 265022, 264693, 52644150, 264691, 33657003, 264693, 33657109, 33657349, 264628, 18108370, 60431528, 18108374, 35686423, 65274791, 60170394, 83375044, 87168518, 22279000, 22279002, 264486, 265006, 264288
3119	86728796 (6237, 6238)		Contains protein domain (PF00328) - Histidine acid phosphatase	264488, 264509, 264510, 264511, 264512, 264288, 264486
3120	87344040 (6239, 6240)	Novel Protein sim. GBank g01501918jgIA037863.1(AE14315, (AE143152) putative NADH oxidoreductase complex subunit [Caenorhabditis elegans]	UNCLASSIFIED	52844507, 52845156, 52646365, 52646842, 22278994, 22278996, 22278997, 22278998, 264269, 29331824, 33657377, 22278999, 52646405, 265008, 52646317, 87168497, 87168559, 21906765, 52644150, 33657003, 18108374, 264637
3121	84110735 (6241, 6242)	Novel Protein sim. GBank g01401877efNF_001088.1(AQR1 - acrosin	UNCLASSIFIED	18108332, 29331822, 29331824, 29331826, 264905, 265007, 55812038, 265019, 18108351, 284682, 264288, 264766, 21906764, 21906765, 21906768, 21906769, 55811957, 18108365, 18108366, 27486265, 18108374, 18108381, 18108394, 2279000, 22279002, 264482
3122	18114528 (6243, 6244)	Novel Protein sim. GBank g02439517 (AC002563) - putative RHO-RAC effector protein, 95% similarity to P49205 (P10g134680) [Homo sapiens]	Contains protein domain (PF00780) - GNI domain	264638
3123	38063503 (6245, 6246)	Novel Protein sim. GBank g01490326jgIA035412.1(AE00171 - (AE001714) oxidoreductase, short chain dehydrogenase/reductase family [Thermotoga maritima]	UNCLASSIFIED	56181686, 264259, 56714117, 60432289, 29331826, 29331827, 264907, 264908, 564328, 265009, 60433356, 33657402, 60433357, 264785, 18108391, 264288, 29344627, 29344628, 33657003, 33657109, 18108392, 56526486
3124	87786599 (6247, 6248)	Novel Protein sim. GBank g01490326jgIA035412.1(AE00171 - (AE001714) oxidoreductase, short chain dehydrogenase/reductase family [Thermotoga maritima]	UNCLASSIFIED	56181686, 264259, 56714117, 60432289, 29331826, 29331827, 264907, 264908, 564328, 265009, 60433356, 33657402, 60433357, 264785, 18108391, 264288, 29344627, 29344628, 33657003, 33657109, 18108392, 56526486
3125	91216607 (6249, 6250)	Novel Protein sim. GBank g01490326jgIA035412.1(AE00171 - (AE001714) oxidoreductase, short chain dehydrogenase/reductase family [Thermotoga maritima]	UNCLASSIFIED	56181686, 264259, 56714117, 60432289, 29331826, 29331827, 264907, 264908, 564328, 265009, 60433356, 33657402, 60433357, 264785, 18108391, 264288, 29344627, 29344628, 33657003, 33657109, 18108392, 56526486

3128	95337205 (6251, 6252)			UNCLASSIFIED	22278996, 264490, 264259, 60432049, 29331822, 60432269, 29146496, 52644045, 56182435, 265009, 60433438, 265010, 87168559, 285017, 265018, 58611130, 294763, 264683, 264369, 264685, 29146829, 33657023, 264693, 33657109, 10108374, 55911576, 18105385, 60432113, 22277002, 35652496, 22276986, 22276989, 28331826, 29331822, 2943438, 87168559, 264694, 29162765, 218108374, 29162769, 264629, 18108374, 18108377, 22279000, 22279002
3127	91639233 (6253, 6254)	Novel Protein sim. GBank gll424220[mimCAA6694, 11- (AL021687) putative protein [Arabidopsis thaliana]			
3128	87674330 (6255, 6256)	Novel Protein sim. GBank gll385528 (AF090133) - lin-7-A [Rattus norvegicus]		Contains protein domain (PF00595) - misc. channel PDZ domain (Also known as DHR or GLOF).	22278996, 264259, 52644045, 265008, 21906754, 265017, 265018, 21906763, 18108376, 18108387, 22279000, 22279002
3129	87755412 (6257, 6258)	Novel Protein sim. GBank gll1335273 (AC003058) - hypothetical protein [Arabidopsis thaliana]		Contains protein domain (PF00400) - kinase WD domain, G-beta repeat	56182575, 264259, 29331825, 23331828, 52644045, 56182435, 60433355, 264600, 264682, 264763, 264764, 264369, 264488, 264686, 55811957, 264682, 33657023, 33657109, 60432113, 264464, 264466, 264636
3130	14939860 (6259, 6260)	Novel Protein sim. GBank gll3239463 (AF064553) - NSD1 protein [Mus musculus]			
3131	95351469 (6261, 6262)	Novel Protein sim. GBank gll1848277 (J08138) - telomerase-associated protein TP-1 [Homo sapiens]		Contains protein domain (PF00400) - WD domain, G-beta repeat	56182575, 264259, 29331824, 264497, 56182435, 264594, 60433438, 55812038, 33105954, 21906734, 33657084, 87168474, 264686, 264687, 264688, 35911957, 265020, 265021, 265022, 265023, 265024, 33657023, 33657109, 33657182, 27486261, 33657349, 65274791, 60170394, 56182323, 83373044, 87168518, 264464

3132	95415459 (6263, 6264)	Novel Protein sim. GBank g1680647/bpA027113 [AF13293 - (AF13293) CGI-03 protein [Homo sapiens]	Contains protein domain (P100789) - ubiquitin UBX domain	52644507, 52648412, 52648355, 65245172, 56182574, 522715994, 22276995, 3556266, 56994075, 222715998, 22276997, 22276998, 22276999, 60432049, 52645080, 29331822, 29331824, 29331825, 60432285, 29331826, 29331827, 29331828, 30960052, 264508, 52644045, 56182435, 264910, 60170831, 60432223, 60433356, 33857402, 59812038, 52646317, 21900794, 52644296, 65688542, 87108559, 265017, 265018, 265019, 264448, 264288, 264305, 52644223, 21900785, 21900786, 21900787, 21900788, 21900789, 26501917, 265020, 33857021, 60170915, 52644159, 264692, 33857023, 52644129, 33857109, 33857182, 27466261, 27466262, 27466265, 33857349, 35695763, 18108374, 18108376, 55811578, 35695855, 18108385, 18108387, 56526486, 87168516, 60432113, 22279002
3133	87379414 (6265, 6266)	Novel Protein sim. GBank g14507613/jrefNP_003738, lipTNKS - TANKYASE	polymerase	22276994, 22276996, 254505, 265006, 265007, 87168559, 264780, 21900767, 18108374, 22279000, 22279002, 264553, 264595, 264369, 264685, 264626, 264596
3134	94843816 (6267, 6268)	Novel Protein sim. GBank g11729827/jpP54633[ALA_DICDI - HOMOLOG]		
3135	86389356 (6269, 6270)	Novel Protein sim. GBank g13093478 (AF012821) - flavinogen-binding protein (Streptococcus esui)	struct	22276998, 264595, 29331826, 33857402, 18108386, 265019, 80433435, 265019, 264764, 264288, 264769, 264659, 265020, 27466262, 263972, 65274791, 264557, 264558
3136	94445839 (6271, 6272)	Novel Protein sim. GBank g1627101/jpP544952 - probable cancer protein c2 - Caenorhabditis elegans	Contains protein domain (P100153) - Mitochondrial carrier proteins	22276995, 22276996, 22276997, 22276998, 22276999, 264259, 29331822, 29331825, 29331826, 29331827, 29331828, 264511, 265008, 21900754, 87168474, 265011, 87168559, 265017, 265018, 265019, 18108351, 254692, 264769, 21900765, 21900766, 21900767, 21900769, 55811957, 35895917, 25020, 265021, 52644150, 18108370, 18108374, 22279000, 22279002, 264482, 264486
3137	88257947 (6273, 6274)	Novel Protein sim. GBank g13342750 (AC005531) - R31341_1 [Homo sapiens]	UNCLASSIFIED	

3144	95353529 (5287, 6298) Novel Protein sim. GBank g1494448(emb CA94332.1) - (AL050225) hypothetical protein [Homo sapiens]			264488, 18108396, 22276986, 35696286, 22277997, 22277999, 28331826, 29331827, 35696052, 29331828, 264106, 265006, 265007, 265009, 33857402, 85585542, 265011, 18108351, 264448, 264389, 2100765, 21900766, 21900767, 2855280, 1905011, 3245416, 2168581, 18108350, 18108349, 18108348, 5683233, 83373044, 22279000, 22277002, 264567
3145	86611857 (6298, 6290) Novel Protein sim. GBank g13876709(emb CA803330) - (281118) Similarity to Human endosomal protein P162 (TR Q15079); cDNA EST EMBL Z11487 comes from this gene; cDNA EST EMBL Z14556 comes from this gene; cDNA EST EMBL D27011 comes from this gene; cDNA EST EMBL D27015 comes from this gene	UNCLASSIFIED		18108397, 29331824, 29146769, 20281100, 365006, 55812038, 265010, 21900766, 29146627, 21905769, 29146784, 264692, 33657023, 33657109, 35695763, 253881, 56182323, 87186518
3146	87756314 (6291, 6292) Novel Protein sim. GBank g12135746(jp S65950) - mitogen inducible gene mlg 2 - human	Contains protein domain (PF00169) - struct PH domain		264259, 29331826, 29331828, 29331830, 264510, 264511, 265007, 265009, 264600, 265017, 18108351, 264448, 264389, 21900766, 265021, 294992, 33657109, 18108374, 35696423, 35695955, 60432113, 264384
3147	94848512 (6293, 6294) Novel Protein sim. GBank g1387427(emb CA807315) 1 - (236259) predicted using Genefinder; cDNA EST YK3184723 comes from this gene; cDNA EST YK3154125 comes from this gene [Caenorhabditis elegans]	Contains protein domain (PF00702) - halacid dehalogenase-like hydrolase	UNCLASSIFIED	56181686, 35696286, 60432040, 264259, 56182101, 29331828, 60432039, 35696052, 60432040, 264558, 264557, 264559, 60433356, 60433358, 265010, 264448, 264288, 265022, 33857023, 33657109, 60431528, 66271791, 294631, 56182323, 264404, 22278002
3148	96382169 (6295, 6296) Novel Protein sim. GBank g1625232(gb AA040851.1) (AF-083108) sirutin type 3 [Homo sapiens]	UNCLASSIFIED		35696286, 35696052, 264511, 85658542, 87168474, 284754, 35696423, 264555, 264558, 264557, 264559, 83373044, 56526486, 60432113
3149	95300548 (6297, 6298) Novel Protein sim. GBank g14205448 (AF-02777) - FYVE finger-containing phospholipase kinase [Mus musculus]	Contains protein domain (PF01363) - FYVE zinc finger	UNCLASSIFIED	29331822, 35696052, 264109, 20148628, 18108381
3150	97655472 (6299, 6300) Novel Protein sim. GBank g1387454(jp BCA71683) - (Y17850) gangliolide-induced differentiation associated protein 1 [Mus musculus]	Contains protein domain (PF00043) - Glutathione S-transferases	UNCLASSIFIED	264259, 29331822, 29331824, 29331825, 29331827, 5264637, 264686, 35695855, 56182323, 264639
3151	87722355 (6301, 6302) Novel Protein sim. GBank g1472391 (M63377) - SFP1 [Saccharomyces cerevisiae]	Contains protein domain (PF00096) - zinc finger, C2H2 type	UNCLASSIFIED	29331822, 265008
3152	86088108 (6303, 6304) Novel Protein sim. GBank g1494448(emb CA94332.1) - (AL050225) hypothetical protein [Homo sapiens]	UNCLASSIFIED		21900764, 87186518, 264605, 21905768, 55841150, 27186264, 35696423, 22279000

3153	95317299 (6305, 6306)	Novel Protein sim. GBank gi4895041 gb AA032705.1 AF 14995 - (AF 14395) oronin 3 [Mus musculus]	Contains protein domain (PF00400) WD domain, G-beta repeat	struct	264488, 52646365, 3568288, 22278596, 22278997, 22278999, 60432048, 264259, 2331326, 61432269, 33656970, 264508, 264803, 33657402, 264595, 6043436, 264804, 264805, 264806, 264807, 264448, 264682, 264764, 264268, 264389, 264758, 21906755, 21906766, 21906767, 21906768, 21906769, 2014874, 265021, 265022, 60170615, 5264150, 264690, 264691, 33657023, 65274620, 33657109, 18108370, 35695655, 264638, 60170394, 87168518, 60432113, 22279000, 22279002, 22278998, 264259, 29331824, 66715502, 265008, 265010, 265017, 18108354, 264691, 33657023, 264693, 20281149, 18108374 264511, 264512, 33657402, 264681, 264683, 33657023, 18108370, 264634, 264639, 18108385, 264595, 264596, 264597, 22278998, 22278999, 60432048, 264259, 29331822, 22278996, 60432049, 264259, 29331822, 29331824, 60714117, 20331825, 29331826, 29331827, 56860692, 52644045, 265007, 265009, 60170831, 60432229, 60433356, 21906754, 33109654, 87168474, 265010, 265017, 265018, 265019, 18108351, 264448, 264288, 264689, 21906766, 21906768, 21906769, 35695617, 265020, 265022, 264692, 18108370, 35696423, 56182323, 22279002
3154	87718573 (6307, 6308)	Novel Protein sim. GBank gi4680661 gb AA027720.1 AF 13294 - (AF 132945) CQI-11 protein [Homo sapiens]	ATPase-associated		
3155	87782394 (6309, 6310)	Novel Protein sim. GBank gi728637 gb F39194 ALU_ HUMAN - III: ALU SUBFAMILY SQ WARNINGS ENTRY III	UNCLASSIFIED		
3156	87737449 (6311, 6312)	Novel Protein sim. GBank gi530606 gb AA045621.1 AC005601 - (AC005601) N- terminal domain myristyltransferase, similar to Q10473 (PDB gi 1709559) [Homo sapiens]	Contains protein domain (PF00642) Similarity to leucin domain of ricin beta-chain, 3 copies.	transferase	
3157	88255577 (6313, 6314)				
3158	80034118 (6315, 6316)	Novel Protein sim. GBank gi530606 gb AA045621.1 AF 15877 - (AF 158778) ASB-3 protein [Homo sapiens]	Contains protein domain (PF00023) - kinase Ank repeat	kinase	
3159	84124114 (6317, 6318)	Novel Protein sim. GBank gi530606 gb AA045621.1 AF 15877 - (AF 158778) ASB-3 (AJ243800) WSC4 homologue [Myriophyllum spicatum]	UNCLASSIFIED		
3160	80221068 (6319, 6320)	Novel Protein sim. GBank gi530606 gb AA045621.1 AF 15877 - (AF 158778) ASB-3 determination protein homolog Fm1a [Mus musculus]	Contains protein domain (PF00023) - struct Ank repeat	struct	56182575, 22278998, 29331824, 264406, 60433356, 264758, 265017, 18108355, 264448, 18108354, 264768, 21906766, 265020, 264691, 264692, 33657109, 18108374, 35696423, 264555, 60170394, 22279000 264559

3161	9807411 (6321, 6322)	284488, 2227895, 2227897, 2227898, 284259, 2833182, 6043228, 2833182, 3284403, 203017, 283816, 284448, 284288, 14866, 2839787, 283830, 10108374, 284488, 284488
------	----------------------	---

Table 2

Tissue ID	Tissue Name	Tissue Information	Disease Association
20281069	192xN	Protein-protein Interactions	Any
20281071	192xN	Protein-protein Interactions	Any
20281149	192xN	Protein-protein Interactions	Any
20281152	192xN	Protein-protein Interactions	Any
264111	276xN	Protein-protein Interactions	Any
264112	276xN	Protein-protein Interactions	Any
263966	384xN	Protein-protein Interactions	Any
263967	384xN	Protein-protein Interactions	Any
264110	552xN	Protein-protein Interactions	Any
18108379	SPH 52.1 (Adrenal Gland)	Adrenal Gland/Suprarenal gland	Adrenoleukodystrophy, Congenital Adrenal Hyperplasia,
18108381	SPH 52.2 (Fetal Lung)	Fetal Lung	Cystic Fibrosis, infection, lung cancer
18108383	SPH 52.3 (B's Lymphoma- Raji)	Lymphoma derived from B cells	Blood cancers, hematopoiesis, leukemia
18108368	SPH 52.5 (Salivary Gland)	Salivary Gland	Dry mouth, infection
18108384	SPH 52.6 (Brain- Thalamus)	Thalamus	Brain cancer, head injury, obesity, neurological disorders, neuropsychiatric disorders
18108394	SPH 53.1 (Adrenal Gland)	Adrenal Gland/Suprarenal gland	Adrenoleukodystrophy, Congenital Adrenal Hyperplasia,
18108355	SPH 53.2 (Fetal Lung)	Fetal Lung	Cystic Fibrosis, infection, lung cancer
18108359	SPH 53.3 (B's Lymphoma- Raji)	Lymphoma derived from B cells	Blood cancers, hematopoiesis, leukemia
18108361	SPH 53.4 (Mammary Gland)	Mammary Gland	Lactation disorders, breast cancer
18108362	SPH 53.5 (Salivary Gland)	Salivary Gland	Dry mouth, infection
18108366	SPH 53.6 (Brain- Thalamus)	Thalamus	Brain cancer, head injury, obesity, neurological disorders, neuropsychiatric disorders
18108354	SPH 54.1 (Adrenal Gland)	Adrenal Gland/Suprarenal gland	Adrenoleukodystrophy, Congenital Adrenal Hyperplasia,
18108392	SPH 54.2 (Fetal Lung)	Fetal Lung	Cystic Fibrosis, infection, lung cancer
18108348	SPH 54.3 (B's Lymphoma- Raji)	Lymphoma derived from B cells	Blood cancers, hematopoiesis, leukemia
18108382	SPH 54.4 (Mammary Gland)	Mammary Gland	Lactation disorders, breast cancer
18108395	SPH 54.5 (Salivary Gland)	Salivary Gland	Dry mouth, infection
18108365	SPH 54.6 (Brain- Thalamus)	Thalamus	Brain cancer, head injury, obesity, neurological disorders, neuropsychiatric disorders
18108397	SPH 55.1 (Adrenal Gland)	Adrenal Gland/Suprarenal gland	Adrenoleukodystrophy, Congenital Adrenal Hyperplasia,
18108398	SPH 55.2 (Fetal Lung)	Fetal Lung	Cystic Fibrosis, infection, lung cancer
18108364	SPH 55.3 (B's Lymphoma- Raji)	Lymphoma derived from B cells	
18108388	SPH 55.4 (Mammary Gland)	Mammary Gland	Lactation disorders, breast cancer
18108358	SPH 55.5 (Salivary Gland)	Salivary Gland	Dry mouth, infection
20281099	SPH 56.2 (MG63)		
20281100	SPH 56.3 (U1SMC)		
264404	SPH 1 (Brain)	Whole Brain	Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, Stroke, Tuberosclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection

264510	SPH.10 (Pancreas)	Pancreas	Pancreatitis, diabetes, pancreatic cancer
264511	SPH.11 (Placenta)	Placenta	Infertility, birth defects
264512	SPH.12 (Thyroid)	Thyroid	Hyperparathyroidism, Hypoparathyroidism
264555	SPH.13 (Bone Marrow)	Bone Marrow	Hemophilia, hypercoagulation, Idiopathic thrombocytopenic purpura, autoimmune disease, allergies, immunodeficiencies, transplantation, Graft versus host,
264556	SPH.14 (Bone Marrow)	Bone Marrow	Hemophilia, hypercoagulation, Idiopathic thrombocytopenic purpura, autoimmune disease, allergies, immunodeficiencies, transplantation, Graft versus host,
264557	SPH.15 (Bone Marrow)	Bone Marrow	Hemophilia, hypercoagulation, Idiopathic thrombocytopenic purpura, autoimmune disease, allergies, immunodeficiencies, transplantation, Graft versus host,
264558	SPH.16 (Bone Marrow)	Bone Marrow	Hemophilia, hypercoagulation, Idiopathic thrombocytopenic purpura, autoimmune disease, allergies, immunodeficiencies, transplantation, Graft versus host,
264559	SPH.17 (Bone Marrow)	Bone Marrow	Hemophilia, hypercoagulation, Idiopathic thrombocytopenic purpura, autoimmune disease, allergies, immunodeficiencies, transplantation, Graft versus host,
264569	SPH.19 (One Fetal tissue and two cell lines)	Mixed	
264687	SPH.19.1 (fetal thymus - CRL7046)	Fetal Thymus	Hemophilia, hypercoagulation, Idiopathic thrombocytopenic purpura, immunodeficiencies
264688	SPH.19.2 (hematopoietic stem cells - CRL2043)	Hematopoietic stem cells	Leukemia, osteoporosis, post-chemotherapeutic stem cell repopulation
264689	SPH.19.3 (osteogenic sarcoma cell lines - HTB96)	Osteogenic Sarcoma	Sarcomas, osteoporosis, osteopetrosis
264690	SPH.19.4 (Fetal Liver)	Fetal liver	Von Hippel-Lindau (VHL) syndrome, Cirrhosis, Transplantation
264691	SPH.19.5 (Heart)	Heart	Cardiomyopathy, Atherosclerosis, Hypertension, Congenital heart defects, Aortic stenosis, Atrial septal defect (ASD), Atrioventricular (A-V) canal defect, Ductus arteriosus, Pulmonary stenosis, Subaortic stenosis, Ventricular septal defect (VSD), valve diseases, Tuberous sclerosis, Scleroderma, Obesity, Transplantation
264692	SPH.19.6 (Spleen)	Spleen	Hemophilia, Hypercoagulation, Idiopathic thrombocytopenic purpura, Immunodeficiencies, Graft versus host
264693	SPH.19.7 (Pituitary)	Pituitary	Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, Stroke, Tuberous sclerosis, hypercalcemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection
264482	SPH.2 (Brain)	Brain	Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, Stroke, Tuberous sclerosis, hypercalcemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection

264600	SPH.21 (Fetal Brain)	Fetal brain	Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, Stroke, Tuberosclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection
264601	SPH.22 (Bone Marrow)	Bone Marrow	Hemophilia, hypercoagulation, Idiopathic thrombocytopenic purpura, autoimmune disease, allergies, immunodeficiencies, transplantation, Graft versus host,
264602	SPH.23 (Thyroid)	Thyroid	Hyperthyroidism and Hypothyroidism
264603	SPH.24 (Pancreas)	Pancreas	Pancreatitis, diabetes, pancreatic cancer
264604	SPH.25 (Lymph Node)	Lymph Node	Lymphedema, Allergies
264605	SPH.26 (Placenta)	Placenta	Infertility, birth defects
264634	SPH.28 (Heart)	Heart	Cardiomyopathy, Atherosclerosis, Hypertension, Congenital heart defects, Aortic stenosis, Atrial septal defect (ASD), Atrioventricular (A-V) canal defect, Ductus arteriosus, Pulmonary stenosis, Subaortic stenosis, Ventricular septal defect (VSD), valve diseases, Tuberosclerosis, Scleroderma, Obesity, Transplantation
264635	SPH.29 (Fetal Kidney)	Fetal Kidney	Diabetes, Autoimmune disease, Renal artery stenosis, Interstitial nephritis, Glomerulonephritis, Polycystic kidney disease, Systemic lupus erythematosus, Renal tubular acidosis, IgA nephropathy, Hypercalcaemia, Lesch-Nyhan syndrome
264483	SPH.3 (Bone Marrow)	Bone marrow	Hemophilia, hypercoagulation, Idiopathic thrombocytopenic purpura, autoimmune disease, allergies, immunodeficiencies, transplantation, Graft versus host,
264636	SPH.30 (Lymph Node)	Lymph Node	Lymphedema, Allergies
264637	SPH.31 (Pancreas)	Pancreas	Pancreatitis, diabetes, pancreatic cancer
264638	SPH.32 (Thyroid)	Thyroid	Hyperthyroidism and Hypothyroidism
264639	SPH.33 (Fetal Brain)	Fetal brain	Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, Stroke, Tuberosclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection
264484	SPH.4 (Bone Marrow)	Bone Marrow	Hemophilia, hypercoagulation, Idiopathic thrombocytopenic purpura, autoimmune disease, allergies, immunodeficiencies, transplantation, Graft versus host,
264758	SPH.44.1 (Kidney)	Kidney	Diabetes, Autoimmune disease, Renal artery stenosis, Interstitial nephritis, Glomerulonephritis, Polycystic kidney disease, Systemic lupus erythematosus, Renal tubular acidosis, IgA nephropathy, Hypercalcaemia, Lesch-Nyhan syndrome
264760	SPH.44.2 (Fetal Liver)	Fetal Liver	Von Hippel-Lindau (VHL) syndrome, Cirrhosis, Transplantation

264762	SPH.44.3 (Heart)	Heart	Cardiomyopathy, Atherosclerosis, Hypertension, Congenital heart defects, Aortic stenosis, Atrial septal defect (ASD), Atrioventricular (A-V) canal defect, Ductus arteriosus, Pulmonary stenosis, Subaortic stenosis, Ventricular septal defect (VSD), valve diseases, Tuberosus sclerosis, Scleroderma, Obesity, Transplantation
264764	SPH.44.4 (Prostate)	Prostate	Prostate Cancer
264766	SPH.44.5 (Spleen)	Spleen	Hemophilia, Hypercoagulation, Idiopathic thrombocytopenic purpura, Immunodeficiencies, Graft versus host
264768	SPH.44.6 (pituitary)	Pituitary	Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, Stroke, Tuberosus sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection
264769	SPH.44.7 (Uterus)	Uterus	Infertility, birth defects
264905	SPH.48.1 (Burkitt's Lymphoma- Raji)	Burkitt's Lymphoma	Lymphoma, blood cancers
264906	SPH.48.2 (Thalamus- Brain)	Thalamus	Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, Stroke, Tuberosus sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection
264907	SPH.48.3 (Adrenal Gland)	Adrenal Gland/Suprarenal gland	Adrenoleukodystrophy, Congenital Adrenal Hyperplasia,
264908	SPH.48.4 (Fetal Lung)	Fetal Lung	Cystic Fibrosis, infection, lung cancer
264909	SPH.48.5 (Salivary Gland)	Salivary Gland	Dry mouth, infection
264910	SPH.48.6 (Mammary Gland)	Mammary Gland	Lactation disorders, breast cancer
265006	SPH.50.1 (B's lymphoma)	Burkitt's Lymphoma	Lymphoma, blood cancers
265007	SPH.50.2 (thalamus)	Thalamus	Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, Stroke, Tuberosus sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection
265008	SPH.50.3 (adrenal gland)	Adrenal Gland/Suprarenal gland	Adrenoleukodystrophy, Congenital Adrenal Hyperplasia,
265009	SPH.50.4 (fetal lung)	Fetal Lung	Cystic Fibrosis, infection, lung cancer
265010	SPH.50.5 (salivary gland)	Salivary Gland	Dry mouth, infection
265011	SPH.50.6 (mammary gland)	Mammary Gland	Lactation disorders, breast cancer
18108385	SPH.51.1 (MCF-7)	Breast Cancer	Breast Cancer
18108370	SPH.51.2 (CCRF-CEM)	Cancer Cell line	Cancer
18108374	SPH.51.3 (K-562)	Cancer Cell line	Cancer
18108351	SPH.51.4 (OVCAR-3)	Ovarian cancer	Ovarian cancer
18108372	SPH.51.5 (HL-60)	Cancer Cell line	Cancer
264486	SPH.6 (Bone Marrow)	Bone Marrow	Hemophilia, hypercoagulation, Idiopathic thrombocytopenic purpura, autoimmune disease, allergies, immunodeficiencies, transplantation, Graft versus host,

264508	SPH.8 (Fetal Brain)		Von Hippel-Lindau (VHL) syndrome , Alzheimer's disease, Stroke, Tuberous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection
264509	SPH.9 (Lymph Node)	Lymph Node	Lymphedema , Allergies
20798451	SRH 56.3 (U/SMC)		
264487	SRH.1 (Brain)	Brain	Von Hippel-Lindau (VHL) syndrome , Alzheimer's disease, Stroke, Tuberous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection
264534	SRH.11 (Bone marrow)	Bone Marrow	Hemophilia, hypercoagulation, Idiopathic thrombocytopenic purpura, autoimmune disease, allergies, immunodeficiencies, transplantation, Graft versus host,
264535	SRH.12 (Bone marrow)	Bone Marrow	Hemophilia, hypercoagulation, Idiopathic thrombocytopenic purpura, autoimmune disease, allergies, immunodeficiencies, transplantation, Graft versus host,
264563	SRH.19 (Fetal Brain)	Fetal brain	Von Hippel-Lindau (VHL) syndrome , Alzheimer's disease, Stroke, Tuberous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection
264488	SRH.2 (Bone Marrow)	Bone Marrow	Hemophilia, hypercoagulation, Idiopathic thrombocytopenic purpura, autoimmune disease, allergies, immunodeficiencies, transplantation, Graft versus host,
264564	SRH.20 (Lymph Node)	Lymph Node	Lymphedema , Allergies
264565	SRH.21 (Pancreas)	Pancreas	Pancreatitis, diabetes, pancreatic cancer
264566	SRH.22 (Placenta)	Placenta	Infertility, birth defects
264567	SRH.23 (Thyroid)	Thyroid	Hyperthyroidism and Hypothyroidism
264591	SRH.25 (Fetal Brain)	Fetal brain	Von Hippel-Lindau (VHL) syndrome , Alzheimer's disease, Stroke, Tuberous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection
264592	SRH.26 (Bone Marrow)	Bone Marrow	Hemophilia, hypercoagulation, Idiopathic thrombocytopenic purpura, autoimmune disease, allergies, immunodeficiencies, transplantation, Graft versus host,
264593	SRH.27 (thyroid)	Thyroid	Hyperthyroidism and Hypothyroidism
264594	SRH.28 (Pancreas)	Pancreas	Pancreatitis, diabetes, pancreatic cancer
264595	SRH.29 (Lymph Node)	Lymph Node	Lymphedema , Allergies
264489	SRH.3 (Bone Marrow)	Bone Marrow	Hemophilia, hypercoagulation, Idiopathic thrombocytopenic purpura, autoimmune disease, allergies, immunodeficiencies, transplantation, Graft versus host,

264596	SRH.30 (Placenta)	Placenta	Infertility, birth defects
264628	SRH.33 (Fetal Kidney)	Fetal kidney	Diabetes, Autoimmune disease, Renal artery stenosis, Interstitial nephritis, Glomerulonephritis, Polycystic kidney disease, Systemic lupus erythematosus, Renal tubular acidosis, IgA nephropathy, Hypercalcemia, Lesch-Nyhan syndrome
264629	SRH.34 (Lymph Node)	Lymph Node	Lymphedema, Allergies
264630	SRH.35 (Pancreas)	Pancreas	Pancreatitis, diabetes, pancreatic cancer
264631	SRH.36 (thyroid)	Thyroid	Hyperthyroidism and Hypothyroidism
264632	SRH.37 (Fetal Brain)	Fetal Brain	Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, Stroke, Tuberous sclerosis, hypercalcemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection
264490	SRH.4 (Bone Marrow)	Bone Marrow	Hemophilia, hypercoagulation, idiopathic thrombocytopenic purpura, autoimmune disease, allergies, immunodeficiencies, transplantation, Graft versus host
264681	SRH.43.1 (fetal thymus - CRL7046)	Fetal Thymus	Hemophilia, hypercoagulation, idiopathic thrombocytopenic purpura, immunodeficiencies
264682	SRH.43.2 (hematopoietic stem cells - CRL2043)	Hematopoietic stem cells	Leukemia, osteoporosis, post-chemotherapy stem cell repopulation
264683	SRH.43.3 (osteogenic sarcoma cell lines - HTB96)	Osteogenic Sarcoma	Sarcomas, osteoporosis, osteopetrosis
264684	SRH.43.4 (Fetal Liver)	Fetal Liver	Von Hippel-Lindau (VHL) syndrome, Cirrhosis, Transplantation
264685	SRH.43.6 (Spleen)	Spleen	Hemophilia, Hypercoagulation, idiopathic thrombocytopenic purpura, Immunodeficiencies, Graft versus host
264686	SRH.43.7 (pituitary)	Pituitary	Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, Stroke, Tuberous sclerosis, hypercalcemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection
264757	SRH.44.1 (Kidney)	Kidney	Diabetes, Autoimmune disease, Renal artery stenosis, Interstitial nephritis, Glomerulonephritis, Polycystic kidney disease, Systemic lupus erythematosus, Renal tubular acidosis, IgA nephropathy, Hypercalcemia, Lesch-Nyhan syndrome
264759	SRH.44.2 (Fetal Liver)	Fetal Liver	Von Hippel-Lindau (VHL) syndrome, Cirrhosis, Transplantation
264761	SRH.44.3 (Heart)	Heart	Cardiomyopathy, Atherosclerosis, Hypertension, Congenital heart defects, Aortic stenosis, Atrial septal defect (ASD), Atrioventricular (A-V) canal defect, Ductus arteriosus, Pulmonary stenosis, Subaortic stenosis, Ventricular septal defect (VSD), valve diseases, Tuberous sclerosis, Scleroderma, Obesity, Transplantation
264763	SRH.44.4 (Prostate)	Prostate	Prostate Cancer
264765	SRH.44.5 (Spleen)	Spleen	Hemophilia, Hypercoagulation, idiopathic thrombocytopenic purpura, immunodeficiencies, Graft versus host

264767	5RH.44.6 (Pituitary)	Pituitary	Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, Stroke, Tuberous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection
264828	5RH.46.1 (Lymph Node)	Lymph Node	Lymphedema, Allergies
264887	5RH.47.5 (Fetal Liver)	Fetal Liver	Von Hippel-Lindau (VHL) syndrome, Cirrhosis, Transplantation
18108377	5RH.50.1 (B's lymphoma)	Burkitt's Lymphoma	Lymphoma, blood cancers
18108380	5RH.50.2 (thalamus)	Thalamus	Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, Stroke, Tuberous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection
18108396	5RH.50.3 (adrenal gland)	Adrenal Gland/Suprarenal gland	Adrenoleukodystrophy, Congenital Adrenal Hyperplasia
18108391	5RH.50.4 (fetal lung)	Fetal Lung	Airway diseases, infection
18108357	5RH.50.5 (salivary gland)	Salivary Gland	Dry mouth, infection
18108390	5RH.50.6 (mammary gland)	Mammary Gland	Lactation disorders, breast cancer
264532	5RH.9 (Bone Marrow)	Bone Marrow	Hemophilia, hypercoagulation, idiopathic thrombocytopenic purpura, autoimmune disease, allergies, immunodeficiencies, transplantation, Graft versus host
263974	736xN		
263976	736xN		
263981	736xN		
20281166	96xN		
20281169	96xN		
20281171	96xN		
263994	cDNA-ORF Selection		
264080	Mx96		
21906754	NQH 6.1 (HH729)		
22278996	NQH 6.10 (PbEC)	Endothelial cells	heart disease, cancer
22278997	NQH 6.11 (CAEC)	Endothelial cells	heart disease, cancer
22278998	NQH 6.12 (CSC)	Cancer Cell line	Cancer
22278999	NQH 6.13 (NHNPC)	Cancer Cell line	Cancer
22279000	NQH 6.14 (NHMC-RM)	Cancer Cell line	Cancer
22279002	NQH 6.15 (Hypothalamus)	Hypothalamus	Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, Stroke, Tuberous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection, Obesity
21906764	NQH 6.2 (In Dated Platelets)	Platelets	Clotting diseases, stroke
21906765	NQH 6.3 (HuVec)	Endothelial cells	heart disease, cancer
87168474	NQH 6.3 (Sized-HUVEC)	Endothelial cells	heart disease, cancer
21906766	NQH 6.4 (UdMVEC-myc)	Cancer Cell line	Cancer
21906767	NQH 6.5 (NHEM-neo)	Cancer Cell line	Cancer
21906768	NQH 6.6 (NHEK)	Cancer Cell line	Cancer
21906769	NQH 6.7 (ByCAEC)	Endothelial cells	heart disease, cancer
22278994	NQH 6.8 (NHA)	Cancer Cell line	Cancer

22278995	NQH 6.9 (PrSC)	Cancer Cell line	Cancer
27486261	NQH 7.1 (Jurkat E6-untreated)	Cancer Cell line	Cancer
27486262	NQH 7.2 (TF1-untreated)	Cancer Cell line	Cancer
27486264	NQH 7.3 (U87-untreated)	Cancer Cell line	Cancer
27486265	NQH 7.4 (THP1-untreated)	Cancer Cell line	Cancer
29331822	NQH 8.1 (Brain-amygdala)		Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, Stroke, Tuberos sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection
29331824	NQH 8.2 (Brain-hippocampus)		Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, Stroke, Tuberos sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection
29331825	NQH 8.3 (Brain-substantia nigra)		Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, Stroke, Tuberos sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection
29331826	NQH 8.4 (small intestine)	Small intestine	digestive diseases, obesity, diabetes
29331827	NQH 8.5 (Spinal cord)	Spinal chord	paralysis, neurodegenerative disorders
29331828	NQH 8.6 (stomach)	Stomach	Stomach cancer
29331830	NQH 8.7 (Trachea)	Trachea	Airway diseases, infection
87168518	NQH 9.1 (Sized-MG-63 treatment pool)		
87168559	NQH 9.2 (Sized-HEPG2 untreated)		
35695763	NQH.10.1 (MCF-7untreated)	Cancer Cell line	Cancer
35695855	NQH.10.2 (U-937 treatment pool)	Cancer Cell line	Cancer
35695917	NQH.10.3 (JAR)	Cancer Cell line	Cancer
35696052	NQH.10.4 (PA-1)	Cancer Cell line	Cancer
35696286	NQH.10.5 (CADMEC)	Endothelial cells	heart disease, cancer
35696423	NQH.10.6 (CADMEC LA)	Endothelial cells	heart disease, cancer
52644045	NQH.11.1 (SK-PN-DW)	Cancer Cell line	Cancer
52644150	NQH.11.2 (Chorionic Villus Cells)	Chorionic villus	fertility, birth defects
52644229	NQH.11.3 (A549)	Cancer Cell line	Cancer
52644256	NQH.11.4 (U265B1)	Cancer Cell line	Cancer
52644332	NQH.11.5 (Daoy)	Cancer Cell line	Cancer
52644507	NQH.11.6 (SW1783)	Cancer Cell line	Cancer
52645080	NQH.12.1 (U-118MG)	Cancer Cell line	Cancer
52645129	NQH.12.2 (A204)	Cancer Cell line	Cancer
52645156	NQH.12.3 (T24)	Cancer Cell line	Cancer
52646317	NQH.12.4 (G-401)	Cancer Cell line	Cancer
52646365	NQH.12.5 (CaSk)	Cancer Cell line	Cancer
52646842	NQH.12.6 (SHP-77)	Cancer Cell line	Cancer

60424179	NQH.14.1 (Yale75_breast carcinoma)	Breast carcinoma	Breast Cancer
60424269	NQH.14.2 (Yale78B_ovarytumor)	Ovary tumor	Ovarian cancer
60431528	NQH.14.3 (Yale79_prostateBPH)	Prostate	Prostate Cancer
60431602	NQH.14.4 (Yale80_ProstateAdenocarcinoma)	Prostate	Prostate Cancer
60431735	NQH.14.5 (Yale86_UterineMyoma)	Uterine Myoma	Uterine Cancer
60431850	NQH.14.6 (Yale207_Myometrium)	Myometrium	Fertility
60432049	NQH.15.1 (Yale99_cervix)	Cervix	Osteoporosis, cervical cancer
60432113	NQH.15.2 (Yale45_spleenITP)		Hemophilia, Hypercoagulation, Idiopathic thrombocytopenic purpura, Immunodeficiencies, Graft versus host
60432229	NQH.15.3 (Yale16_Skin)	Skin	wound healing, melanoma
60432289	NQH.15.4 (Yale137_Parotid)		
60433356	NQH.15.5 (Yale38_SmallIntestine)	Small intestine	digestive diseases, obesity, diabetes
60433438	NQH.15.6 (Yale28_ColonAscending)	Colon	Colon cancer
65274444	NQH.17.1 (Larynx)	Larynx	Cancer
65274572	NQH.17.2 (Duodenum)	Duodenum	
65274620	NQH.17.3 (Kidney, Primary tumors)		Diabetes, Autoimmune disease, Renal artery stenosis, Interstitial nephritis, Glomerulonephritis, Polycystic kidney disease, Systemic lupus erythematosus, Renal tubular acidosis, IgA nephropathy, Hypercalcemia, Lesch-Nyhan syndrome
65274727	NQH.17.4 (Lung Pleura, normal)	Lung	Airway diseases, infection
65274791	NQH.17.5 (Lung, Normal Adult)	Lung	Airway diseases, infection
83373044	NQH.18.230 (Pooled adrenal gland, placenta)	Adrenal Gland/Suprarenal gland	Adrenoleukodystrophy, Congenital Adrenal Hyperplasia,
85658542	NQH.18.560 (Pooled uterus, BeWo pool)	Uterus	Infertility, birth defects
33656970	NQH.9.1 (MG-63_treatment pool)	Cancer Cell line	Cancer
33657023	NQH.9.2 (HEPG2 untreated)		Von Hippel-Lindau (VHL) syndrome, Cirrhosis, Transplantation
33657084	NQH.9.3 (PC3_untreated)	Cancer Cell line	Cancer
33657109	NQH.9.4 (TF-1_TPA)	Cancer Cell line	Cancer
33657182	NQH.9.5 (TF-1_TPO)	Cancer Cell line	Cancer
33657349	NQH.9.6 (TF-1_Hemin)	Cancer Cell line	Cancer
33657402	NQH.9.7 (HFDPC)	Cancer Cell line	Cancer
264259	NQH1 (Mixture of eight adult & two fetal tissues)		
264288	NQH2 (Ten tissues plus lymphocyte control)		
264448	NQH3 (Bone Marrow)	Bone Marrow	Hemophilia, hypercoagulation, Idiopathic thrombocytopenic purpura, autoimmune disease, allergies, immunodeficiencies, transplantation, Graft versus host,
265017	NQH4.1 (lymph node)	Lymph Node	Lymphedema, Allergies

265018	NQH4.2 (fetal kidney)	Fetal Kidney	Diabetes, Autoimmune disease, Renal artery stenosis, Interstitial nephritis, Glomerulonephritis, Polycystic kidney disease, Systemic lupus erythematosus, Renal tubular acidosis, IgA nephropathy, Hypercalcaemia, Lesch-Nyhan syndrome
66712502	NQH4.2 (Sized)		
265019	NQH4.3 (pituitary gland)		Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, Stroke, Tuberosus sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection, Obesity
66714117	NQH4.3 (Sized)		
265020	NQH4.4 (testis)	testis	Infertility, birth defects
265021	NQH4.5 (fetal liver)	Fetal Liver	Von Hippel-Lindau (VHL) syndrome, Cirrhosis, Transplantation
265022	NQH4.6 (thyroid)	Thyroid	Hyperthyroidism and Hypothyroidism
18108376	NQH5.1 (MCF-7)	Breast cancer	Breast Cancer
18108387	NQH5.2 (CCRF-CEM)	Cancer Cell line	Cancer
264952	NRL1: HPLC FRACTIONATION OF RE-LIG		
263971	Old BB3 Baits		
263969	Old BB5 Baits		
263975	ORFSEL		
263972	OTHER Baits		
263978	pGALORF		
264106	PPBAITS		
264088	QC-YA7		
264089	QC-YA8		
264102	Reassorted Interactors		
264369	RRH.1		
60170394	RRH.10.1 (MCF-7untreated)	Breast cancer	Breast Cancer
60170615	RRH.10.2 (U-937_treatment pool)	Cancer Cell line	Cancer
60170831	RRH.10.3 (JAR)	Cancer Cell line	Cancer
60174639	RRH.11.8 (HeLa)	Cancer Cell line	Cancer
264113	mQEA Baits		
263973	RRQEA_B5 baits		
29146498	SRD 3.1 (SKMC)	Cancer Cell line	Cancer
29146499	SRD 3.2 (SKMC)	Cancer Cell line	Cancer
29147620	SRD 3.3 (RPTEC)	Cancer Cell line	Cancer
29148627	SRD 3.4 (HRCE)	Cancer Cell line	Cancer
29148629	SRD 3.6 (HRE)	Cancer Cell line	Cancer
29148784	SRD 3.7 (HRE)	Cancer Cell line	Cancer
55810764	SRD.7.1 (Lymph Node)	Lymph Node	Lymphedema, Allergies
55811150	SRD.7.2 (pancreas)	Pancreas	Pancreatitis, diabetes, pancreatic cancer
55811386	SRD.7.3 (Adrenal Gland)	Adrenal Gland/Suprarenal gland	Adrenoleukodystrophy, Congenital Adrenal Hyperplasia,

55811576	SRD.7.4 (Pituitary Gland)	Pituitary	Von Hippel-Lindau (VHL) syndrome , Alzheimer's disease, Stroke, Tuberous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection, Obesity
55811957	SRD.7.5 (Fetal Liver)	Fetal Liver	Von Hippel-Lindau (VHL) syndrome, Cirrhosis, Transplantation
55812038	SRD.7.6 (Fetal Kidney)	Fetal kidney	Diabetes, Autoimmune disease, Renal artery stenosis, Interstitial nephritis, Glomerulonephritis, Polycystic kidney disease, Systemic lupus erythematosus, Renal tubular acidosis, IgA nephropathy, Hypercalcaemia, Lesch-Nyhan syndrome
56181562	SRD.8.1 (Lymph Node)	Lymph Node	Lymphedema , Allergies
56181686	SRD.8.2 (Pancreas)	Pancreas	Pancreatitis, diabetes, pancreatic cancer
56182181	SRD.8.3 (Adrenal Gland)	Adrenal Gland/Suprarenal gland	Adrenoleukodystrophy , Congenital Adrenal Hyperplasia,
56182323	SRD.8.4 (Pituitary Gland)	Pituitary	Von Hippel-Lindau (VHL) syndrome , Alzheimer's disease, Stroke, Tuberous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection, Obesity
56182435	SRD.8.5 (Fetal Liver)	Fetal Liver	Von Hippel-Lindau (VHL) syndrome, Cirrhosis, Transplantation
56182575	SRD.8.6 (Fetal Kidney)		Diabetes, Autoimmune disease, Renal artery stenosis, Interstitial nephritis, Glomerulonephritis, Polycystic kidney disease, Systemic lupus erythematosus, Renal tubular acidosis, IgA nephropathy, Hypercalcaemia, Lesch-Nyhan syndrome
32833986	SRD4: HL adapter		
56526486	SRD5: 1 arr fragments		
33109954	SRD5: long-RXRJ		
56994075	SRD9:1 (CS/SC)	Cancer Cell line	Cancer
263977	TSC Screen 1		

Disclosed Sequences of ORFX Nucleic Acid and Polypeptide Sequences

<210> 1
 <211> 312
 <212> DNA
 <213> Homo sapiens

<400> 1
 gtcgacagtg tcattggggcgc ccgaccggcg atcggcgttg cgagcacgtt cgggtaccgcg
 60
 agagggaatcg gtaccgcgcg acgaggagga cgggcgggca gctcggcgac gagtgggtccg
 120
 gcgagcnncg atccgtctgc ctccccacca ccgatctcgg cagccaggtt gtcgaggatg
 180
 tcgctgggtgg agcgcgtgctc cgcgtctgct tcttgctcag ccgtcttgcg ggactgggccc
 240
 tgctctctcgg ccttggtcag ggcagcgtca atgctggcga aatccgattc tgccttcggc
 300
 tcagagacgc gt
 312

<210> 2
 <211> 104
 <212> PRT
 <213> Homo sapiens

<400> 2
 Val Asp Ser Val Met Gly Ala Arg Pro Ala Ile Gly Val Ala Ser Thr
 1 5 10 15
 Phe Gly Thr Ala Arg Gly Ile Gly Thr Ala Arg Arg Gly Gly Arg Ala
 20 25 30
 Gly Ser Ser Ala Thr Ser Gly Pro Ala Ser Xaa Asp Pro Ser Ala Ser
 35 40 45
 Pro Pro Pro Ile Ser Ala Ala Arg Leu Ser Arg Met Ser Leu Val Glu
 50 55 60
 Arg Cys Ser Ala Ser Ala Ser Cys Ser Ala Val Leu Arg Asp Trp Ala
 65 70 75 80
 Cys Ser Ser Ala Leu Val Arg Ala Ala Ser Met Ser Ala Lys Ser Asp
 85 90 95
 Ser Ala Phe Gly Ser Glu Thr Arg
 100

<210> 3

<211> 987

<212> DNA

<213> Homo sapiens

<400> 3

gagctcggta cccccgtcgt cgccgtcaag gcgcaggtca aagtgggtgg ccgtggcaag
 60
 gctggcgggtg taaagatcgc taagtccccg gccgaggcgg ctgaacgcgc cgaggccatc
 120
 ctcgggtatgg acatcaaggg tcacaccgtt cacaagggtga tggctcgtga aggtgctgac
 180
 attgccgagg aatactactt ctcgatcttg ctcgaccgtg gagagcgtcg ctaccttgcg
 240
 atgtgctcgc gtgagggtgg catggacatc gagacccttg ctaaggagcg ccccgagggt
 300
 ctcgccaaagg tgccgggtcga cccgattgac ggcggttgacg atgctaaagc ccgcgaaatc
 360
 cttagcgagg cagggttccc cgactctgag caggacgcta tcgtcccggc tggttctcaag
 420
 ctgtgggaga cctaccgtga cgaggatgcc accctcgtcg aggtcaaaccc gatgatcaag
 480
 accggcgatg gacgcatact ggctatcgac ggcaagatga ctggtgacaa caacgcattc
 540
 ttccgccagc ctgaccgcgc cggccttggtg gatcgcgcga ccaccgaccc gctcgagttg
 600
 cgtgccggcg agctcgggtc caactacgtc aagcttgacg gcaacgtcgg cgtcatcgga
 660
 aacgggtgctg gtctgggtcat gagcaccctt gactgcgttg cgtacgcggg tgagaaacttc
 720
 ccgggatctc cagctccgcg taacttcctc gacatcggtg gcggcgctc gcccgagatc
 780
 atggccaaag gccttgacct catcatgagt gacgagcagg tgcgttcggt gtctgtcaac
 840
 gtcttttggt gtatcacgcg ctgtgaccag gtggcgcgtt gtatcaaggg cgctctcgag
 900
 aagttggggc acaaggccgt gaagccgctc gtcgtccgtc tggacggcaa cgctgtggcc
 960
 gagggcagaa agattctcga ggaattc
 987

<210> 4

<211> 329

<212> PRT

<213> Homo sapiens

<400> 4

Glu	Leu	Gly	Thr	Pro	Val	Val	Ala	Val	Lys	Ala	Gln	Val	Lys	Val	Gly
1				5					10				15		
Gly	Arg	Gly	Lys	Ala	Gly	Gly	Val	Lys	Ile	Ala	Lys	Ser	Pro	Ala	Glu
			20					25				30			
Ala	Ala	Glu	Arg	Ala	Glu	Ala	Ile	Leu	Gly	Met	Asp	Ile	Lys	Gly	His
		35					40					45			

```

Thr Val His Lys Val Met Val Ala Glu Gly Ala Asp Ile Ala Glu Glu
50          55          60
Tyr Tyr Phe Ser Ile Leu Leu Asp Arg Gly Glu Arg Arg Tyr Leu Ala
65          70          75          80
Met Cys Ser Arg Glu Gly Gly Met Asp Ile Glu Thr Leu Ala Lys Glu
85          90          95
Arg Pro Glu Ala Leu Ala Lys Val Pro Val Asp Pro Ile Asp Gly Val
100         105         110
Asp Asp Ala Lys Ala Arg Glu Ile Leu Ser Glu Ala Gly Phe Pro Asp
115         120         125
Ser Glu Gln Asp Ala Ile Val Pro Ala Val Leu Lys Leu Trp Glu Thr
130         135         140
Tyr Arg Asp Glu Asp Ala Thr Leu Val Glu Val Asn Pro Met Ile Lys
145         150         155         160
Thr Gly Asp Gly Arg Ile Leu Ala Ile Asp Gly Lys Met Thr Val Asp
165         170         175
Asn Asn Ala Ser Phe Arg Gln Pro Asp Arg Ala Gly Leu Val Asp Arg
180         185         190
Ala Thr Thr Asp Pro Leu Glu Leu Arg Ala Gly Glu Leu Gly Leu Asn
195         200         205
Tyr Val Lys Leu Asp Gly Asn Val Gly Val Ile Gly Asn Gly Ala Gly
210         215         220
Leu Val Met Ser Thr Leu Asp Cys Val Ala Tyr Ala Gly Glu Asn Phe
225         230         235         240
Pro Gly Ser Pro Ala Pro Ala Asn Phe Leu Asp Ile Gly Gly Gly Ala
245         250         255
Ser Ala Glu Ile Met Ala Asn Gly Leu Asp Leu Ile Met Ser Asp Glu
260         265         270
Gln Val Arg Ser Val Phe Val Asn Val Phe Gly Gly Ile Thr Ala Cys
275         280         285
Asp Gln Val Ala Leu Gly Ile Lys Gly Ala Leu Glu Lys Leu Gly Asp
290         295         300
Lys Ala Val Lys Pro Leu Val Val Arg Leu Asp Gly Asn Ala Val Ala
305         310         315         320
Glu Gly Arg Lys Ile Leu Glu Glu Phe
325

```

<210> 5

<211> 622

<212> DNA

<213> Homo sapiens

<400> 5

```

acgcgtgggc tagacctgaa tccattctct aaaaacagc ctccctcctt gccatgggat
60
tccctgggggt ttccacagcg cacagctcta atggctctgca gcagggttacc ttgttcccca
120
gaacatagctt tgctcataaca tctctgcagg gttctcccaa acccctttctt gectggcaac
180
agctgacatc acacctagct gtaagtcctt gtagatcgca aattactttt tgggactggg
240
gggtgacagg ggcattgggg taatagcctt ctageccctt ttgaggggaa cacatgggtg
300

```

aggcattttt ggggctggga agtggggggc tgggtgtccc tggatggctg tgctggcctc
 360
 tggctgcaag ggagaggggc acaggcaagg acatgacccc cgtcaaccct gagcccccctc
 420
 cagaaattta accagagcct gtcctcctct tcttgctgc cccaacatc tcacaatccc
 480
 tccgtgtgat gcagatgtct ccatctactc tacagacacc tgcaactatc attcccttga
 540
 tccgtggtaa ttaggagggg actcctctgt gaagaaccgc ttctaccatc cttcttttaga
 600
 aactctttct ccactgggat cc
 622

<210> 6
 <211> 121
 <212> PRT
 <213> Homo sapiens

<400> 6
 Met Ser Leu Pro Val Pro Leu Ser Leu Ala Ala Arg Gly Gln His Ser
 1 5 10 15
 His Pro Gly Asp Thr Arg Pro Pro Leu Pro Ser Pro Lys Ile Ala Ser
 20 25 30
 Pro Met Cys Phe Pro Gln Lys Gly Leu Glu Gly Tyr Tyr Pro Asn Ala
 35 40 45
 Pro Ala Thr Pro Ser Leu Gln Lys Val Ile Cys Asp Leu Gln Gly Leu
 50 55 60
 Thr Ala Arg Cys Asp Val Ser Cys Cys Gln Ala Glu Arg Gly Leu Gly
 65 70 75 80
 Glu Pro Cys Arg Asp Val Met Thr Ser Tyr Val Leu Gly Asn Lys Val
 85 90 95
 Thr Cys Cys Arg Pro Leu Glu Leu Trp Pro Val Lys Thr Pro Gly Asn
 100 105 110
 Pro Met Ala Arg Arg Glu Thr Val Leu
 115 120

<210> 7
 <211> 480
 <212> DNA
 <213> Homo sapiens

<400> 7
 agatcttttc agaagcgggtg ggtggggcgt tcatatccac gtgaatgggc ccaaaaggca
 60
 agaaaaacca cactctctct cccctcccca gtggtgtgtg ctttccttagg gacaatagga
 120
 tgaatgggct ttcagtgtgg ggacagcaaa acatgcacta gggcccagag tggcagttct
 180
 cttggtgtgg agagtgcctg ccacaggcct tggccagagc cgtgagggga gtggtgtgtg
 240
 aaaggccacc tccacgtggg taagcgtgag gacttggact tctctggcac tgagatggga
 300
 cctcctgcct gtgggagtca tctggccacc accctggggc cagtaaaggt tggagctaga
 360

agggctgctcc tccctgactt gagctctgag ggctttgcct gccagccag agcggaagg
 420
 cacaggggac cctcggggac gcccatggcc accctgggga agacagggct cctcacgcgt
 480

<210> 8
 <211> 119
 <212> PRT
 <213> Homo sapiens

<400> 8
 Met Gly Phe Gln Cys Gly Asp Ser Lys Thr Cys Thr Arg Ala Gln Ser
 1 5 10 15
 Gly Ser Ser Leu Gly Val Glu Ser Ala Cys His Arg Pro Trp Pro Glu
 20 25 30
 Pro Val Arg Glu Trp Cys Val Lys Gly His Leu His Val Gly Lys Arg
 35 40 45
 Glu Asp Leu Asp Phe Ser Gly Thr Glu Met Gly Pro Ala Cys Gly
 50 55 60
 Ser His Leu Ala Thr Thr Leu Gly Pro Val Lys Val Gly Ala Arg Arg
 65 70 75 80
 Val Val Leu Pro Asp Leu Ser Ser Glu Gly Phe Ala Cys Pro Ala Arg
 85 90 95
 Ala Ala Arg His Arg Gly Pro Ser Gly Thr Pro Met Ala Thr Leu Gly
 100 105 110
 Lys Thr Gly Leu Leu Thr Arg
 115

<210> 9
 <211> 428
 <212> DNA
 <213> Homo sapiens

<400> 9
 tccggaaatt ttccatgaga gactggggca tgcagcagaa gtggatgtct gttctcctgc
 60
 ctctgctgct accttacaat gatccgttct tccccctctc ctctcctggc aacagctggc
 120
 tcccagggat gctggatgac ctctttcagt ccatgttctc gtgcgccctg ctgctcttct
 180
 ggctgtgcgt gtaccacggg attcgtgtcc agggagaaag aaagtgttta actttctatt
 240
 tgccctaaatt cttcattgtt ggactattgt ggttggtctc tggtacgcta ggaatatggc
 300
 aaacagtgtaa cgaattacat gatccaatgt accagtatcg agttgatacc ggaatttttc
 360
 aggggaatgaa ggtctctctc atgggtggtg cagcgggtga cattctgtac ctctgttctc
 420
 tgatagtg
 428

<210> 10
 <211> 138
 <212> PRT

<213> Homo sapiens

<400> 10

```
Met Arg Asp Trp Gly Ile Glu Gln Lys Trp Met Ser Val Leu Leu Pro
 1           5           10           15
Leu Leu Leu Leu Tyr Asn Asp Pro Phe Pro Leu Ser Phe Leu Val
 20           25           30
Asn Ser Trp Leu Pro Gly Met Leu Asp Asp Leu Phe Gln Ser Met Phe
 35           40           45
Leu Cys Ala Leu Leu Leu Phe Trp Leu Cys Val Tyr His Gly Ile Arg
 50           55           60
Val Gln Gly Glu Arg Lys Cys Leu Thr Phe Tyr Leu Pro Lys Phe Phe
 65           70           75           80
Ile Val Gly Leu Leu Trp Leu Ala Ser Val Thr Leu Gly Ile Trp Gln
 85           90           95
Thr Val Asn Glu Leu His Asp Pro Met Tyr Gln Tyr Arg Val Asp Thr
100           105           110
Gly Asn Phe Gln Gly Met Lys Val Phe Phe Met Val Val Ala Ala Val
115           120           125
Tyr Ile Leu Tyr Leu Leu Phe Leu Ile Val
130           135
```

<210> 11

<211> 453

<212> DNA

<213> Homo sapiens

<400> 11

```
cttaagaatc gcctcactca acggtcagct tgccgacccat gcccgccctga taatgccccg
 60
gaatgttgca ggcaaaactca agaccggtgg ccttggtgaa ggtccagggtc agctcgggcg
120
acttgccccg ctgcaccagc acgtgtgttg ggtcgtcatg cttcatgccc cccatatacg
180
catgccccat ggccggtgtg tccatcttgc ccatgccggt ggccgtgagc atgccgctgg
240
cttgcatctt gagcatttct ttctggtgtt cggcgtgcat gcgccatca cccagattga
300
atcgtgcag taactggcct ttgttgacca gcacaaagcg caccggtctca ccggtttta
360
catccagagc cttggggcga aaggaaatgt cctgcagggt gacttccacg gtgcgcgtgg
420
ctttatcgcc cggtgcccgg tgcccaaacg cgt
453
```

<210> 12

<211> 130

<212> PRT

<213> Homo sapiens

<400> 12

```
Met Leu Gln Ala Asn Ser Arg Pro Val Ala Leu Val Lys Val Gln Val
 1           5           10           15
Ser Ser Ala Asp Leu Pro Gly Ser Thr Ser Thr Leu Leu Gly Ser Ser
```


[illegible]

```
<210> 13
<211> 2034
<212> DNA
<213> Homo sapiens
```

```

400> 13
nagcgcttgc gcgtagtccc ctctctgcc tccgtcctga gtcctctgct gcccgctgctg
60
ggcggtggcca agcaggagac ggtgcgcgtg gccttctgct ccggggacct gccgctctctc
120
tggggccctac catgctggca tttctctcca tgtgtcaaac acatgggttc agccagcgaa
180
gattccatgg gacctctctg tgtgggacgt gtgctcccca ccacaaatg aacgtctctt
240
gtttgatcat ggagggggtg gtggtctctg tggctggagc agcctggggc cagaggaagc
300
cgtatacaac ggetctgcag cgcttcagcg aggggtgcct ggagtagccta gccaacctgg
360
accgagcccc agaccceacg gtcaggaagg acgcctttgc caccgacatc ttcagcgctt
420
acgatgttct ctcccatcag tggctgcaga gtcgagaagc caagctccgt cttgccgttg
480
tggaggtctt ggggcctatg agccatctgc tgcccagtga gaggctggaa gagcagctgc
540
ccaagctcct ccttgggatt ctgcacctct acaagaagca cgcagagacc ttctacttgt
600
ccaagagcct gggccagatc ctcgaggcag ctgtgagtgt gggcagccgc aactctggaga
660
ccagcctgga tgccctcttg gctgcactgc actcccagat ctgtgtgcct gtggagctct
720
caagccccct ggtgatgagt aaccagaagg aggtgctgcg ctgcttcact gtgctggcct
780
gcagctgcgc tgaccgccta ctggccttcc tgctgcccag gctggacac accaatgaga
840
ggaccgcgct gggcaccctg cagggggcca aacatgtcat caactcaact cgtgctcaaa
900
tgggaagataa aaagcccttt atcctgtctt ccatgaggct tcctctcctg aacaccaaca
960

```

gcaaggtgaa gcgggcagtg gtgcaggtga ttagcgccca tggcccacca cggctacctg
 1020
 gagcagcctg gaggtgagge gatgatcgag taaatcgtgc agcagtgctgc getgcccccc
 1080
 gagcaggagc ctgagaagcc agggccccgc agcaaggacc ccaaggccga cagcgtgcgg
 1140
 gccatcagcg tgcgcaccct ctacctgtgc agcaccaccg tggacaggat gagtcactgc
 1200
 ctctggccat acctgtctca gttcctcacc cctgtgctgc tctactggggc cctgactccg
 1260
 ctctgcagga gcctcgtgca tctggcgagc aagaggcagg aggccggggc cgacgccttc
 1320
 ctcacccagt acgacgcccc tgcgagcctc cgtctccct atgctgtaac cggaagactg
 1380
 ttggtttgtg cttccagccc ctacctaggg gacggagctg gggcagcggc gctgcgcctc
 1440
 ctacagtgtc tgcacccaaa cattcaccct ttgctgggtc agcattggga aacgactgtc
 1500
 ccgctgctgc tggggtaacct ggatgagcac acagaagaga ccctgccaca ggaggagtgg
 1560
 gaggagaagc tgttgatggt gagggccggg gtacggccca tcctgggcct taagggtgtg
 1620
 tctggcctgg ggggtgctgg ggtggcagag gctggggcgc ctgcctcgac ctcacctcgt
 1680
 ggtttggctg gggagccaag gatcaggcag catcaaggct gaagacccca gcagccttgc
 1740
 agcggggggc ttgctgtgac aaggcacccg ccctctagca gtgcagccc caagcgtcgg
 1800
 gggcaacctc tcacctctgc tggtagacca actgtggcat ggctgtcccc tgaggggttg
 1860
 ctctgcccgc ccggcctccc gctggaaggc ggtctgcagc ccctgcagcc acagcacatg
 1920
 gggatgtgcc caggctccag ccagccctgt gaggggtcgg gctccagccc cctcagtggc
 1980
 atcttgccct gcagttcctg cgagaccccc tggccatcat ttctgacaac gcgt
 2034

<210> 14

<211> 222

<212> PRT

<213> Homo sapiens

<400> 14

Ile	Val	Gln	Gln	Cys	Ala	Leu	Pro	Pro	Glu	Gln	Glu	Pro	Glu	Lys	Pro
1				5					10					15	
Gly	Pro	Gly	Ser	Lys	Asp	Pro	Lys	Ala	Asp	Ser	Val	Arg	Ala	Ile	Ser
			20						25				30		
Val	Arg	Thr	Leu	Tyr	Leu	Val	Ser	Thr	Thr	Val	Asp	Arg	Met	Ser	His
		35					40					45			
Val	Leu	Trp	Pro	Tyr	Leu	Leu	Gln	Phe	Leu	Thr	Pro	Val	Arg	Phe	Thr
		50				55					60				
Gly	Ala	Leu	Thr	Pro	Leu	Cys	Arg	Ser	Leu	Val	His	Leu	Ala	Gln	Lys
65					70				75					80	
Arg	Gln	Glu	Ala	Gly	Ala	Asp	Ala	Phe	Leu	Ile	Gln	Tyr	Asp	Ala	His

			85					90					95		
Ala	Ser	Leu	Pro	Ser	Pro	Tyr	Ala	Val	Thr	Gly	Arg	Leu	Leu	Val	Val
			100					105					110		
Ser	Ser	Ser	Pro	Tyr	Leu	Gly	Asp	Gly	Arg	Gly	Ala	Ala	Ala	Leu	Arg
			115				120					125			
Leu	Leu	Ser	Val	Leu	His	Pro	Asn	Ile	His	Pro	Leu	Leu	Gly	Gln	His
			130			135						140			
Trp	Glu	Thr	Thr	Val	Pro	Leu	Leu	Leu	Gly	Tyr	Leu	Asp	Glu	His	Thr
			145		150					155				160	
Glu	Glu	Thr	Leu	Pro	Gln	Glu	Glu	Trp	Glu	Glu	Lys	Leu	Leu	Met	Val
			165					170					175		
Arg	Ala	Gly	Val	Arg	Pro	Ile	Leu	Gly	Leu	Lys	Val	Leu	Ser	Gly	Leu
			180				185					190			
Gly	Gly	Ala	Gly	Val	Ala	Glu	Ala	Gly	Pro	Pro	Ala	Ser	Thr	Ser	Pro
			195			200					205				
Arg	Gly	Leu	Ala	Gly	Glu	Pro	Arg	Ile	Arg	Gln	His	Gln	Gly		
			210			215				220					

```
<210> 15
<211> 363
<212> DNA
<213> Homo sapiens
```

```

400> 15
naccgcgttgc tggctcgcca cggaagggc catgtcggct gcgatattcg caagccggcg
60
gtgggttcga tcttgcctc gtgctggaac aacccgatca tggaccggcg gttggtccg
120
ttgcaggaca ccaatgacac ctctatggcc aacatgcaga agaaccgtac ctattcgatc
180
atcccgcgta tcgcgcggcg cgagatcacc ccggacaac tgatcgccct cggcgcggtg
240
cggaagaaat acgatctgta caccaaagat accgcggccc agcggatcga cctgttcg
300
gccagttgc acgaattgcc gcagatctg ggcgagctgg tggatgcgg attcgagac
360
ggt
363

```

```
<210> 16
<211> 121
<212> PRT
<213> Homo sapiens
```

Xaa																Lys																Gly																Val																Gly																Cys																Asp																Ile																																																																																																																																																																																																															
1																5																10																15																20																25																30																35																40																45																50																55																60																																																																																																																															
Xaa																Lys																Gly																Val																Gly																Cys																Asp																Ile																Pro																Ala																Leu																Val																Gln																Thr																Ser																Tyr																Ile																Leu																Gly																Ala															
1																5																10																15																20																25																30																35																40																45																50																55																60																																																																																																																															
Xaa																Lys																Gly																Val																Gly																Cys																Asp																Ile																Pro																Ala																Leu																Val																Gln																Thr																Ser																Tyr																Ile																Leu																Gly																Ala															
1																5																10																15																20																25																30																35																40																45																50																55																60																																																																																																																															
Xaa																Lys																Gly																Val																Gly																Cys																Asp																Ile																Pro																Ala																Leu																Val																Gln																Thr																Ser																Tyr																Ile																Leu																Gly																Ala															
1																5																10																15																20																25																30																35																40																45																50																55																60																																																																																																																															
Xaa																Lys																Gly																Val																Gly																Cys																Asp																Ile																Pro																Ala																Leu																Val																Gln																Thr																Ser																Tyr																Ile																Leu																Gly																Ala															
1																5																10																15																20																25																30																35																40																45																50																55																60																																																																																																																															
Xaa																Lys																Gly																Val																Gly																Cys																Asp																Ile																Pro																Ala																Leu																Val																Gln																Thr																Ser																Tyr																Ile																Leu																Gly																Ala															
1																5																10																15																20																25																30																35																40																45																50																55																60																																																																																																																															
Xaa																Lys																Gly																Val																Gly																Cys																Asp																Ile																Pro																Ala																Leu																Val																Gln																Thr																Ser																Tyr																Ile																Leu																Gly																Ala															
1																5																10																15																20																25																30																35																40																45																50																55																60																																																																																																																															
Xaa																Lys																Gly																Val																Gly																Cys																Asp																Ile																Pro																Ala																Leu																Val																Gln																Thr																Ser																Tyr																Ile																Leu																Gly																Ala															
1																5																10																15																20																25																30																35																40																45																50																55																60																																																																																																																															
Xaa																Lys																Gly																Val																Gly																Cys																Asp																Ile																Pro																Ala																Leu																Val																Gln																Thr																Ser																Tyr																Ile																Leu																Gly																Ala															
1																5																10																15																20																25																30																35																40																45																50																55																60																																																																																																																															
Xaa																Lys																Gly																Val																Gly																Cys																Asp																Ile																Pro																Ala																Leu																Val																Gln																Thr																Ser																Tyr																Ile																Leu																Gly																Ala															
1																5																10																15																20																25																30																35																40																45																50																55																60																																																																																																																															
Xaa																Lys																Gly																Val																Gly																Cys																Asp																Ile																Pro																Ala																Leu																Val																Gln																Thr																Ser																Tyr																Ile																																																															

```

65              70              75              80
Ala Lys Lys Tyr Asp Leu Tyr Thr Lys Ile Thr Gly Gly Gln Arg Ile
      85              90              95
Asp Leu Phe Gly Ala Gln Leu His Glu Leu Pro Gln Ile Trp Gly Glu
      100             105             110
Leu Val Asp Ala Gly Phe Glu Thr Gly
      115              120

```

```

<210> 17
<211> 682
<212> DNA
<213> Homo sapiens

```

```

<400> 17
gaattccatt ttgtggagta agaggtgact ggggtatagg gtacaaccca tagccatcca
60
tgttcatctt tgttttgaat ataattggct agaagatata catatatcta tgtaacttcc
120
tctagcatcc tocagtatgg aggctgcatt aagactgcat gaaggagagg gagagaaggg
180
agaaacagag cagctggaca agaggacagg tatagggaat aagggagagg ccagtaaggc
240
aggaaagacc ctccgtgaca aaggggcagg gaacagaact caaacattta atggcaggta
300
accagggtta gaatggtaaa ttgaaagggt aatataaaagg gagaatgggt aatgaattt
360
tctgaaatta attgctgtgt ttatagtttt tagccatgca tcggaatcac ctccaggactc
420
cactccaat caattatata tctgggggag gaccaaggcg ttggtatttt tcagaagctc
480
cactgggtgat tctgacagca cagctaggat taagaaactg atcaatggga acagcatgcc
540
tgttgcagag gagcttcctt gggaaatgct acacacagaa catcaatctt ccttccccac
600
tcttgagatc cctcattctt tggcaccagg aacagttgca attagtaaac cctggttccc
660
tgctgtctca caaatgcaa ga
682

```

```

<210> 18
<211> 110
<212> PRT
<213> Homo sapiens

```

```

<400> 18
Met Asn Phe Leu Lys Leu Ile Ala Val Phe Ile Val Phe Ser His Ala
1              5              10              15
Ser Glu Ser Pro Gln Asp Ser Thr Pro Asn Gln Leu Tyr Ile Trp Gly
      20              25              30
Arg Thr Lys Ala Leu Val Phe Phe Arg Ser Ser Thr Gly Asp Ser Asp
      35              40              45
Ser Thr Ala Arg Ile Lys Lys Leu Ile Asn Gly Asn Ser Met Pro Val
      50              55              60
Ala Glu Glu Leu Pro Trp Glu Met Ser His Thr Glu His Gln Ser Ser

```

```

65              70              75              80
Phe Pro Thr Pro Glu Ile Pro His Ser Leu Ala Pro Gly Thr Val Ala
            85              90              95
Ile Ser Lys Pro Trp Phe Pro Ala Val Ser Gln Ile Ala Arg
            100              105              110

```

<210> 19

<211> 515

<212> DNA

<213> Homo sapiens

<400> 19

```

cttggctggc agacatggga cctgcttccc tcttacaccc cagtcttggc aaggatcatg
60
cccccatctc aactatgta gccagctctg ctgttcactt agtcactaca gtttgcctct
120
cgctctgcagt gcagctcttgg gctataagaa acactgggcc actcaatacc tcccccttt
180
tgccctctct cctcctctgg tccatgggtg ggggtggggg gagcccagtt tcagcaccag
240
cagctgggagc ccataaccaca ctcatctttc agttctggct gtggggagccc ctcccacagg
300
tttcagttcc ccaagcccca ggcctgagtt tttttattg caaagctggg ttgtgtgtgt
360
ggctagctcc caggcgtgtg aggtgcagct tgctaagtaa gagctaggaa agagaatagg
420
gtctctgtgt aggtgtccag tctgaaggaa tgccctgggat acttctctca gcagttcctt
480
ctcacagtct cctggctgct ccgcatgtca gatct
515

```

<210> 20

<211> 130

<212> PPT

<213> Homo sapiens

<400> 20

```

Met Gly Pro Ala Ser Leu Leu His Pro Ser Leu Gly Lys Asp His Ala
1              5              10              15
Pro Ile Ser Thr Met Leu Ala Ser Leu Ala Val His Leu Val Thr Thr
20              25              30
Val Cys Phe Ser Ser Ala Val Gln Ser Trp Ala Ile Arg Asn Thr Gly
35              40              45
Pro Leu Asn Thr Ser Pro Leu Leu Ala Leu Leu Leu Trp Ser Met
50              55              60
Gly Gly Val Gly Gly Ser Pro Val Ser Ala Pro Ala Ala Gly Ala His
65              70              75              80
Thr Thr Leu Ile Phe Gln Phe Trp Leu Trp Glu Pro Leu Pro Gln Val
85              90              95
Ser Val Pro Gln Ala Pro Gly Leu Ser Phe Phe Tyr Cys Lys Ser Trp
100              105              110
Leu Leu Leu Trp Leu Ala Pro Arg Arg Val Arg Cys Ser Leu Leu Ser
115              120              125
Lys Ser

```

130

<210> 21
 <211> 390
 <212> DNA
 <213> Homo sapiens

<400> 21
 gtgcgcacaa aagagcacgt tcgcaagggg aggaagagcg tgccaccggt tctgccgagc
 60
 tagacgcggg gcctatgggt gcggaggacc atggagtga gcgagtaaga ctatgatgatg
 120
 caacaaatgt gcctgagggt gaaatggcac gagccagtgc caatgagggc atgacacctg
 180
 ttaaccacga caaataccct tctgtccttt taaatgaagc ggcccaggct tcattactgg
 240
 atacaatgac tgcttgacct gatgggttca caattgagca attggagctt acacgatctc
 300
 tatgttatga aagagtatta gcacatcgat cctcatggga tcgttcagcc ctgggtcaag
 360
 aattaaagca agttgtccaa ggcattccatn
 390

<210> 22
 <211> 105
 <212> PRT
 <213> Homo sapiens

<400> 22
 Met Val Ala Glu Asp His Gly Val Lys Arg Val Arg Leu Asp Asp Ala
 1 5 10 15
 Thr Asn Val Pro Glu Gly Glu Met Ala Arg Ala Ser Ala Asn Glu Gly
 20 25 30
 Met Thr Pro Val Asn His Asp Lys Tyr Pro Ser Val Leu Leu Asn Glu
 35 40 45
 Ala Ala Gln Ala Ser Leu Leu Asp Thr Met Thr Ala Cys Thr Asp Gly
 50 55 60
 Phe Thr Ile Glu Gln Leu Glu Leu Thr Arg Ser Leu Cys Tyr Glu Arg
 65 70 75 80
 Val Leu Ala His Arg Ser Ser Trp Asp Arg Ser Ala Leu Ala Gln Glu
 85 90 95
 Leu Lys Gln Val Val Gln Gly Ile His
 100 105

<210> 23
 <211> 385
 <212> DNA
 <213> Homo sapiens

<400> 23
 ntctcggagg ccgacagcct ggcgggctgg aagccctcgg tgtaccacgt gctgctcctc
 60
 ctgggcctgt tcgcctgct gctgtctctgc tgcgcctcgg ccattgtacac cagcgtggag
 120

ggctgggact acgtgggactc gctctacttc tgettegtca ccttcagcac catcggttc
 180
 ggggacctgg tgagcagcca gcacgcgcgc taccggaacc aggggctcta cgcctgggc
 240
 aacttcctct tcatcctgct cggcgtgtgc tgcatttact cgctcttcaa cgtcatctcc
 300
 atcctcatca agcaggtgct caactggatg ctgcgcaagc tgagctgccg ctgctgcgcg
 360
 cgctgctgcc cggctcctgg cgcgc
 385

<210> 24
 <211> 128
 <212> PRT
 <213> Homo sapiens

<400> 24
 Xaa Ser Glu Ala Asp Ser Leu Ala Gly Trp Lys Pro Ser Val Tyr His
 1 5 10 15
 Val Leu Leu Ile Leu Gly Leu Phe Ala Val Leu Leu Ser Cys Cys Ala
 20 25 30
 Ser Ala Met Tyr Thr Ser Val Glu Gly Trp Asp Tyr Val Asp Ser Leu
 35 40 45
 Tyr Phe Cys Phe Val Thr Phe Ser Thr Ile Gly Phe Gly Asp Leu Val
 50 55 60
 Ser Ser Gln His Ala Ala Tyr Arg Asn Gln Gly Leu Tyr Arg Leu Gly
 65 70 75 80
 Asn Phe Leu Phe Ile Leu Leu Gly Val Cys Cys Ile Tyr Ser Leu Phe
 85 90 95
 Asn Val Ile Ser Ile Leu Ile Lys Gln Val Leu Asn Trp Met Leu Arg
 100 105 110
 Lys Leu Ser Cys Arg Cys Cys Ala Arg Cys Cys Pro Ala Pro Gly Ala
 115 120 125

<210> 25
 <211> 337
 <212> DNA
 <213> Homo sapiens

<400> 25
 ccattgggaga gaccgtgcat tttcttctag gtctgcgtgg gaagtcactg cagagtttctg
 60
 agggaggggag ttccacgctc tgtatttttg aagggtcagt cttgttgctt ggaccagtga
 120
 ggagccccgt gggatccaga ctcgagtggg tggagccggg gcaggtggga gcagagacac
 180
 tggaggaaag ctggtcgaat gcactgtgta ttggaggcca gaaccagcag agggctcctct
 240
 ggggttgatg tagggcaaaa gaaaaagaag gcaccaagcc tggggctctg gttttctctc
 300
 ttacacttgc tgggtggacg gtggtgccac tgaatga
 337

<210> 26

<211> 111
 <212> PRT
 <213> Homo sapiens

<400> 26
 Met Gly Glu Thr Val His Phe Leu Leu Gly Leu Arg Gly Lys Ser Leu
 1 5 10 15
 Gln Ser Phe Glu Glu Gly Ser Ser Gln Leu Cys Ile Phe Glu Gly Ser
 20 25 30
 Val Leu Leu Leu Gly Pro Val Arg Ser Pro Val Gly Ser Arg Leu Glu
 35 40 45
 Trp Val Glu Pro Gly Gln Val Gly Ala Glu Thr Leu Glu Glu Ser Trp
 50 55 60
 Ser Asn Ala Leu Cys Ile Trp Arg Gln Asn Gln Gln Arg Val Leu Trp
 65 70 75 80
 Val Glu Cys Arg Ala Lys Glu Lys Glu Gly Thr Lys Pro Gly Val Trp
 85 90 95
 Val Phe Ser Leu Thr Leu Ala Gly Trp Thr Val Val Pro Leu Asn
 100 105 110

<210> 27
 <211> 333
 <212> DNA
 <213> Homo sapiens

<400> 27
 ccgacgtcga ataccatgc agccgcgcgcg aggatggaga gagcgatgga gcaactcaac
 60
 gcgcctgacgc gctcgtctgcg ccgcgcgcgcg accgtggagt tgcccaggga taatgaaact
 120
 gctgtttata cattaatgcc aatgggtatg gctgatcaac acagggtctgt ttctgaaacta
 180
 ctatcaaat caaaatttga tgtcaattat gcattcggac gtgtgaaaag aagcttgctt
 240
 cacattgcag caaattgtgg atcggtggaa tgcttggttt tgctgttaaa gaaaaggagca
 300
 aatcctaact atcaagatat ttcaggctgt aca
 333

<210> 28
 <211> 111
 <212> PRT
 <213> Homo sapiens

<400> 28
 Pro Thr Ser Asn Ile His Ala Ala Pro Arg Met Glu Arg Ala Met
 1 5 10 15
 Glu Gln Leu Asn Arg Leu Thr Arg Ser Leu Arg Arg Ala Arg Thr Val
 20 25 30
 Glu Leu Pro Glu Asp Asn Glu Thr Ala Val Tyr Thr Leu Met Pro Met
 35 40 45
 Val Met Ala Asp Gln His Arg Ser Val Ser Glu Leu Leu Ser Asn Ser
 50 55 60
 Lys Phe Asp Val Asn Tyr Ala Phe Gly Arg Val Lys Arg Ser Leu Leu


```

65              70              75              80
His Ile Ala Ala Asn Cys Gly Ser Val Glu Cys Leu Val Leu Leu Leu
              85              90              95
Lys Lys Gly Ala Asn Pro Asn Tyr Gln Asp Ile Ser Gly Cys Thr
              100              105              110

```

```

<210> 29
<211> 375
<212> DNA
<213> Homo sapiens

```

```

<400> 29
ncgccgtccg tgctggctat tatgacggcg ggtagcgacc agggcgagga ggtcaactcg
60
gagagctatt tgagcgccgt gacgccgctg agtcccaaag agattcgtca gctgccccgc
120
tacaatatca cgatcaagcg cgtcgtgaac atgacgggca agggccgcac gccgagctgg
180
tactcgctcg tcgtggctgg caatggctcg ggccctcgtg gctatggcga agggcaaatg
240
actaacatca gccgcgcgaa caaaaaggcg ttccacgccg cggtgaaaaa catggacttg
300
gtatcggtcc accggtcgaa gaggggcgcc aacacgctcg agccccccgt cgagggcgccg
360
tgggggcgcta cgcgt
375

```

```

<210> 30
<211> 125
<212> PRT
<213> Homo sapiens

```

```

<400> 30
Xaa Pro Ser Val Leu Ala Ile Met Thr Ala Gly Ser Asp Gln Gly Glu
1          5          10          15
Glu Val Asn Ser Glu Ser Tyr Leu Ser Ala Val Thr Pro Leu Ser Pro
20         25         30
Lys Glu Ile Arg Gln Leu Pro Arg Tyr Asn Ile Thr Ile Lys Arg Val
35         40         45
Val Asn Met Thr Gly Lys Gly Arg Thr Pro Ser Trp Tyr Ser Leu Val
50         55         60
Val Ala Gly Asn Gly Arg Gly Leu Val Gly Tyr Gly Glu Gly Lys Asp
65         70         75         80
Thr Asn Ile Ser Arg Ala Asn Lys Lys Ala Phe His Ala Ala Val Lys
85         90         95
Asn Met Asp Leu Val Ser Val His Arg Ser Lys Ser Gly Ala Asn Thr
100        105        110
Leu Glu Pro Pro Val Glu Gly Arg Trp Gly Ala Thr Arg
115        120        125

```

```

<210> 31
<211> 375
<212> DNA
<213> Homo sapiens

```

<400> 31
 accggtcttg gccctcagctt tgctctgaaa ttgaagtcgg tgccaaaagt ggggaagagc
 60
 gggagcaggc acttacgagc ctgcgcgtca gggatgcttc ctgggccctt gagagtgcag
 120
 agattcctgg atccagagct gcggctgggc ggctgcagct gcgcctggga gtgcagggtt
 180
 cccgcctgc cagctcaaaa ggaatgggg gctcctgcct gttcctggct cctgttggcc
 240
 ctgcagagtg cacaacctta gcgcgccttc ctccactgca gcttacgtct ttgcagcagc
 300
 cactccgat gggctgccac tgccatctgt gagaccataa tgtgtgcaat ttgagactca
 360
 tggcctgcat tgttt
 375

<210> 32
 <211> 118
 <212> PRT
 <213> Homo sapiens

<400> 32
 Met Gln Ala Met Ser Leu Lys Leu His Thr Leu Trp Ser His Arg Trp
 1 5 10 15
 Gln Trp Gln Pro Ile Gly Ser Gly Cys Cys Lys Asp Val Ser Cys Ser
 20 25 30
 Gly Gly Ser Ala Ala Arg Phe Val His Ser Ala Gly Pro Thr Gly Ala
 35 40 45
 Arg Asn Arg Gln Glu Pro Pro Phe Pro Phe Glu Leu Ala Gly Arg Glu
 50 55 60
 Pro Cys Thr Pro Arg Arg Ser Cys Ser Arg Pro Ala Ala Ala Leu Asp
 65 70 75 80
 Pro Gly Ile Ser Ala Leu Ser Gly Ala Gln Glu Ala Ser Leu Thr Arg
 85 90 95
 Arg Leu Val Ser Ala Cys Ser Arg Ser Ser Pro Leu Leu Ala Pro Thr
 100 105 110
 Ser Ile Ser Glu Gln Ser
 115

<210> 33
 <211> 351
 <212> DNA
 <213> Homo sapiens

<400> 33
 ccatgcagcc caaccgttgg cgataaagtc cgtttaggcg ataccaattt atgggcaacc
 60
 attgaacaag atttattaac caaagggtgat gagtgtaaat ttgggtggcgg taaagtgtg
 120
 cggtatggta tggcgcaaag cggcaccgca actcgcgaca atccaaatgt attggatttt
 180
 gtgattacca atgtgatgat cattgatgcc aaattaggca ttatcaaagc cgatattggt
 240

attcgcgatg gtcgtattgt cgggtatcgga caagcaggta accctgacac catgggatgac
 300
 gtacagccaa acatgattat cgggtgctagc acagaagtac ataacgggtgc a
 351

<210> 34
 <211> 117
 <212> PRT
 <213> Homo sapiens

<400> 34
 Pro Cys Ser Pro Thr Val Gly Asp Lys Val Arg Leu Gly Asp Thr Asn
 1 5 10 15
 Leu Trp Ala Thr Ile Glu Gln Asp Leu Leu Thr Lys Gly Asp Glu Cys
 20 25 30
 Lys Phe Gly Gly Gly Lys Ser Val Arg Asp Gly Met Ala Gln Ser Gly
 35 40 45
 Thr Ala Thr Arg Asp Asn Pro Asn Val Leu Asp Phe Val Ile Thr Asn
 50 55 60
 Val Met Ile Ile Asp Ala Lys Leu Gly Ile Ile Lys Ala Asp Ile Gly
 65 70 75 80
 Ile Arg Asp Gly Arg Ile Val Gly Ile Gly Gln Ala Gly Asn Pro Asp
 85 90 95
 Thr Met Asp Asp Val Thr Pro Asn Met Ile Ile Gly Ala Ser Thr Glu
 100 105 110
 Val His Asn Gly Ala
 115

<210> 35
 <211> 355
 <212> DNA
 <213> Homo sapiens

<400> 35
 nngctagctg caccaccacc tgttcatgca ggcagagcgg ccaccctca tggagaaga
 60
 ggaatccact gtattgggca caggcttccct gctggacett ggcaagcagg tgcttggtg
 120
 gtaccaggaa gtccagcgtg tacctcagtg cgtcctcccg ataagtcctc tccaccacat
 180
 ggaacacactg gcccaacagg gtggggggctg ttgctcaaaa ggggtggatac agggcgccgga
 240
 gagtgcctctg cacacagtcc tccactggct caggctccat ggctcggcgc cgggcccgcgt
 300
 ccgacgcttg gtccggggggg cgggggcgggg cgcgccaccg cctcccttca cgcgt
 355

<210> 36
 <211> 118
 <212> PRT
 <213> Homo sapiens

<400> 36
 Xaa Leu Ala Ala Pro Pro Val His Ala Gly Arg Ala Ala Thr Pro

```

      1           5           10           15
His Gly Arg Arg Gly Ile His Cys Ile Gly His Arg Leu Pro Ala Gly
      20           25           30
Pro Trp Gln Ala Gly Ala Trp Leu Val Pro Gly Ser Pro Ala Cys Thr
      35           40           45
Ser Val Arg Pro Pro Asp Lys Ser Ser Pro Pro Gly Thr Pro Gly
      50           55           60
Pro Thr Gly Trp Gly Leu Leu Pro Gln Arg Val Asp Thr Gly Arg Arg
      65           70           75           80
Glu Cys Ser Ala His Ser Pro Pro Leu Ala Gln Ala Pro Trp Leu Gly
      85           90           95
Ala Gly Pro Arg Pro Thr Leu Gly Arg Ala Gly Gly Ala Gly Arg Ala
      100          105          110
Thr Ala Ser Leu His Ala
      115

```

<210> 37
 <211> 492
 <212> DNA
 <213> Homo sapiens

```

<400> 37
acgcgtggcc ttcgtctgcc accaggacgc actcagcccc accgggtttc cggaccgcgc
60
gcaaccatga caaggcgcat gttgtgatct ggttggtatc cttctccgac atgctcgagg
120
gatcggatct ctcggcggtg gtcacggtgc ttgcgcaggc cggctatcgc ccacgggtcc
180
tcgccgacga cgtctgctgc gggttgacgt ggatcactac cggtcagctc gacggtgctc
240
ggcgctgggt gcgcgctggt ctcgacgtgc tggcacccct gtcagacgcc agcgtcccag
300
tcgttggggt agagcgtcc tgcactacgc tctggcgtga tgacgcactc cgcctcctgc
360
cagatgatcc gcgcgtccac cgggtagcca gaaacatgca taccgtcgcc gagatgcttg
420
aggcagcaca gtggacccca cctcgtcgtg caggccacac cctcgtcgtc cagccccatt
480
gtcatcccgcc gg
492

```

<210> 38
 <211> 127
 <212> PRT
 <213> Homo sapiens

```

<400> 38
Met Leu Glu Gly Ser Asp Leu Ser Ala Val Val Thr Val Leu Ala Glu
      1           5           10           15
Ala Gly Tyr Arg Pro Arg Val Leu Ala Asp Asp Val Cys Cys Gly Leu
      20           25           30
Thr Trp Ile Thr Thr Gly Gln Leu Asp Gly Ala Arg Arg Arg Leu Arg
      35           40           45
Ala Gly Leu Asp Val Leu Ala Pro Leu Ser Asp Ala Ser Val Pro Val

```

```

      50              55              60
Val Gly Leu Glu Pro Ser Cys Thr Thr Val Trp Arg Asp Asp Ala Leu
65              70              75              80
Arg Leu Leu Pro Asp Asp Pro Arg Val His Arg Val Ala Arg Asn Met
      85              90              95
His Thr Val Ala Glu Met Leu Glu Ala Ala Gln Trp Thr Pro Pro Ser
      100              105              110
Leu Ala Gly His Thr Leu Val Ala Gln Pro His Cys His Pro Ala
      115              120              125

```

<210> 39

<211> 412

<212> DNA

<213> Homo sapiens

<400> 39

```

aacgaaggtn ccgtacgcgc tctgaaagcc ctgcgtaaag agcgttccga tcgccgggaa
60
gtgatgngca ccgccaaaat gcagggtggtc gaagccgcga gttcaggcaa gattgtcttt
120
gaaatggaag acgtttatta cagcattgcc gaaaaacaac tggtagagcaa cttctctgcg
180
caagtcatgc gtggtgataa aattgcgctg attggcccca acggttgctgg taaaaacgaag
240
ttgctgaaac tgatgttaag taagattcag gcagacagcg gccgtgttca ctgcggtact
300
aaactggaag ttgcgtactt cgaccagcac cgtgctgagc tggatcctga gcgtacgggtg
360
atggataacc tggccgaagg taagcaggaa gtgatggtaa atggccgtgt an
412

```

<210> 40

<211> 137

<212> PRT

<213> Homo sapiens

<400> 40

```

Asn Glu Gly Xaa Val Arg Ala Leu Lys Ala Leu Arg Lys Glu Arg Ser
1              5              10              15
Asp Arg Arg Glu Val Met Xaa Thr Ala Lys Met Gln Val Val Glu Ala
      20              25              30
Ala Ser Ser Gly Lys Ile Val Phe Glu Met Glu Asp Val Tyr Ser
      35              40              45
Ile Ala Gly Lys Gln Leu Val Ser Asn Phe Ser Ala Gln Val Met Arg
      50              55              60
Gly Asp Lys Ile Ala Leu Ile Gly Pro Asn Gly Cys Gly Lys Thr Thr
      65              70              75              80
Leu Leu Lys Leu Met Leu Ser Lys Ile Gln Ala Asp Ser Gly Arg Val
      85              90              95
His Cys Gly Thr Lys Leu Glu Val Ala Tyr Phe Asp Gln His Arg Ala
      100              105              110
Glu Leu Asp Pro Glu Arg Thr Val Met Asp Asn Leu Ala Glu Gly Lys
      115              120              125
Gln Glu Val Met Val Asn Gly Arg Val

```

130

135

<210> 41
 <211> 1080
 <212> DNA
 <213> Homo sapiens

<400> 41
 gaattcaagt ggacacaggc tccacgcccg cgtctcaccg ataagagcta caagcacaaac
 60
 tactatgacg agcgggttct gctcgaagag cgtcttgagc gcactgtggc taaggatttc
 120
 gtcacgacgg aggtcgagcc catgtgggat gcggtgatg tcatgcggat gggtaaggat
 180
 ctcttcaccc agcacggtct gacgacaaat cgggaagtaa tggagtgggt taagcggtac
 240
 taccccgatt tccgcgttca cgcggtgaat ttccctgggg atccgtaccc gatccatcc
 300
 gacgcgacct ttgtgccgct tcgtccgggg ctcatcatca acaaccgaa tcgtccactg
 360
 ccgcaggagc agaggaagat ctccgaggcc aatgactggc agatcgttga tgctgctcag
 420
 ccggcgcaag acacgcctcc agaattgtgc tactcgtctg tgtggctatc aatgaactgc
 480
 ttggtacttg atccgaagac ggtcatctgc gaggcttcgg aagttcatca gatggagcag
 540
 atggacaagc tgggtatgaa cgtcatcccg gtgcctctcc gtgacgcgta cccattcggg
 600
 gggaggctcc actgcgccac agctgatgta tatcgcgaag gtacctgtga ggactacttc
 660
 ccgaatcagg tcgacgaccc gacctgggtg tgagaaaacc ccgtgggtcat gtcagtactg
 720
 acggatctcg gtggtcgggt acggaactta cgttgtccgt taccgggccg ccgggtctga
 780
 tatggcagta tcacgcctag caaaaaggag catgtcatgg acatggagcc gggcatcatc
 840
 aacgtcaaac aggaagtccc aggcgtcggg acgatgaacc agaaagtggg attcgtgtcc
 900
 atgcttcttt ctgcaacggg tatgggggtg gtgggtactt tcgggcgtct cagcactcct
 960
 gtggatccca cgacgggcag taagtacatc atcgggtgatt ttttggccac tggtaggatg
 1020
 atagtcgggg tcctgggatt tctgcttatt atcgtcatca ttgaaaaatg gtctgagctc
 1080

<210> 42
 <211> 230
 <212> PRT
 <213> Homo sapiens

<400> 42
 Glu Phe Lys Trp Thr Gln Ala Pro Arg Pro Arg Leu Thr Asp Lys Ser
 1 5 10 15
 Tyr Lys His Asn Tyr Tyr Asp Glu Arg Val Ser Leu Glu Glu Arg Leu

```

      20      25      30
Glu Arg Thr Val Ala Lys Asp Phe Val Thr Thr Glu Val Glu Pro Met
  35      40      45
Trp Asp Ala Ala Asp Val Met Arg Met Gly Lys Asp Leu Phe Ile Gln
  50      55      60
His Gly Leu Thr Thr Asn Arg Lys Ser Met Glu Trp Phe Lys Arg Tyr
  65      70      75      80
Tyr Pro Asp Phe Arg Val His Ala Val Asn Phe Pro Gly Asp Pro Tyr
      85      90      95
Pro Ile His Ile Asp Ala Thr Phe Val Pro Leu Arg Pro Gly Leu Ile
      100      105      110
Ile Asn Asn Pro Asn Arg Pro Leu Pro Gln Glu Gln Arg Lys Ile Phe
      115      120      125
Glu Ala Asn Asp Trp Gln Ile Val Asp Ala Ala Gln Pro Ala His Asp
      130      135      140
Thr Pro Pro Glu Leu Cys Tyr Ser Ser Val Trp Leu Ser Met Asn Cys
      145      150      155      160
Leu Val Leu Asp Pro Lys Thr Val Ile Cys Glu Ala Ser Glu Val His
      165      170      175
Gln Met Glu Gln Met Asp Lys Leu Gly Met Asn Val Ile Pro Val Ala
      180      185      190
Phe Arg Asp Ala Tyr Pro Phe Gly Gly Gly Leu His Cys Ala Thr Ala
      195      200      205
Asp Val Tyr Arg Glu Gly Thr Cys Glu Asp Tyr Phe Pro Asn Gln Val
      210      215      220
Asp Asp Pro Thr Leu Val
225      230

```

```

<210> 43
<211> 358
<212> DNA
<213> Homo sapiens

```

```

<400> 43
gggcccccca catagtggac acaggtttctt gggatgtcag catggagtgc caagaggtgg
60
gtgaccacct ggtggggaat aaggcgcttc tgggacatag aggctgcctt ccagctgcgc
120
ctggcagagc tgttgacaca acagcatggt ctgcagtgcc gggccactgc cagcgcaccc
180
gatgtccttt aaggatggat ttgggttttc ggattcgctt ggcctatcag cgggagtcct
240
agatcctgaa ggaagtgcag agcccagagg ggaatgatctc gctgaggagc acagctgcct
300
ccctccgcct tgagagagac acaaggcagt tgccactgct caccagtgcc ctgcacgn
358

```

```

<210> 44
<211> 105
<212> PRT
<213> Homo sapiens

```

```

<400> 44
Met Glu Cys Gln Glu Val Gly Asp His Leu Val Gly Asn Lys Ala Leu

```

1	5	10	15
Leu Gly His Arg Gly Cys Leu Pro Ala Ala Pro Gly Arg Ala Val Asp	20	25	30
Thr Thr Ala Trp Ser Ala Val Pro Gly His Cys His Ala His Arg Cys	35	40	45
Pro Leu Arg Met Asp Leu Gly Phe Arg Ile Arg Val Ala Tyr Gln Arg	50	55	60
Glu Ser Gln Ile Leu Lys Glu Val Gln Ser Pro Glu Gly Met Ile Ser	65	70	75
Leu Arg Asp Thr Ala Ala Ser Leu Arg Leu Glu Arg Asp Thr Arg Gln	85	90	95
Leu Pro Leu Leu Thr Ser Ala Leu His	100	105	

<210> 45

<211> 905

<212> DNA

<213> Homo sapiens

<400> 45

gtcgacgata aaggagtatt tgcgcagcag cagtatgatg ctctcggtga ggccgggtttc
 60
 gcggctcctg gaatcccaga gcagtatggt ggcgcagggt cggatcgcat tgcgtccgca
 120
 ataatcatgg aagaggtcgc tcgagttctgt gcgtcgctgt ccaccgtcat atcgccaat
 180
 gagcttggta ccgtccctct cctcaaatac ggtagcggagg agcagaggaa acgttatctt
 240
 tctgaagttg cttegggtaa ggcacttttc ggatatgcgc totccgaggc tgaatgctgga
 300
 tcagatccag ctgcacttaa gtgtcgagcc gacgaagatg gggacagttt cgtcctgaat
 360
 ggcgttaagg cttgggtcac ggaggctggc gaggccaagt acctggtgat atttgcggtt
 420
 actgacccag acgatccgcg ccacagaatc agcgcggtga tgggtccatgc agatgacccg
 480
 ggcattagct acggggctcc ggagcacaaa atgggggatac gcgggtcagt taccagggaa
 540
 gtgggttttca agaatacgcg tatcccacaa gaacagagtaa ttggccctgc agggcacggt
 600
 ctgattgttg ctctaggtac gcttgataac tctcgtgtct cgattgctgc tcaagcagtg
 660
 ggaattgccc aaggagcttt agacattgcc acggattacg tccagaagcg caagcagttt
 720
 ggccagccac tgtccaattt tgagggaatc cagttcatgc tcgcagacat ggcaatgcgt
 780
 ttggaggcgg cgcgagcgct gacatactct gcagctgata gtagtggggc ccagactgac
 840
 gatgtgagtt acctcggcgc ggccggccaaa tgtttcgctt ccgacacagc gatggcagtg
 900
 tgcac
 905

<210> 46

<211> 301

<212> PRT

<213> Homo sapiens

<400> 46

```

Val Asp Asp Lys Gly Val Phe Ala Gln Gln Gln Tyr Asp Ala Leu Val
 1          5          10          15
Glu Ala Gly Phe Ala Ala Pro Gly Ile Pro Glu Gln Tyr Gly Gly Asp
 20          25          30
Gly Ala Asp Ala Ile Ala Ser Ala Ile Ile Met Glu Glu Val Ala Arg
 35          40          45
Val Cys Ala Ser Ser Ser Thr Val Ile Ser Ser Asn Glu Leu Gly Thr
 50          55          60
Val Pro Leu Leu Lys Tyr Gly Ser Glu Glu Gln Arg Lys Arg Tyr Leu
 65          70          75
Ser Glu Val Ala Ser Gly Lys Ala Leu Phe Gly Tyr Ala Leu Ser Glu
 85          90          95
Ala Asp Ala Gly Ser Asp Pro Ala Ala Leu Lys Cys Arg Ala Asp Glu
100          105          110
Asp Gly Asp Ser Phe Val Leu Asn Gly Val Lys Ala Trp Val Thr Glu
115          120          125
Ala Gly Glu Ala Lys Tyr Leu Val Ile Phe Ala Val Thr Asp Pro Asp
130          135          140
Asp Pro Arg His Arg Ile Ser Ala Leu Met Val His Ala Asp Asp Pro
145          150          155
Gly Ile Ser Tyr Gly Ala Pro Glu His Lys Met Gly Ile Arg Gly Ser
165          170          175
Val Thr Arg Glu Val Val Phe Lys Asn Thr Arg Ile Pro Lys Glu Arg
180          185          190
Val Ile Gly Arg Arg Gly His Gly Leu Ser Val Ala Leu Gly Thr Leu
195          200          205
Asp Asn Ser Arg Val Ser Ile Ala Ala Gln Ala Val Gly Ile Ala Gln
210          215          220
Gly Ala Leu Asp Ile Ala Thr Asp Tyr Val Gln Lys Arg Lys Gln Phe
225          230          235
Gly Gln Pro Leu Ser Asn Phe Glu Gly Ile Gln Phe Met Leu Ala Asp
245          250          255
Met Ala Met Arg Leu Glu Ala Ala Arg Ala Leu Thr Tyr Ser Ala Ala
260          265          270
Asp Arg Ser Gly Arg Gln Thr Asp Asp Val Ser Tyr Phe Gly Ala Ala
275          280          285
Ala Lys Cys Phe Ala Ser Asp Thr Ala Met Ala Val Cys
290          295          300

```

<210> 47

<211> 379

<212> DNA

<213> Homo sapiens

<400> 47

```

aagcttgtag agctagtccg aagcggactg tcggtacgcc aagctgctaa aagatgtggg
60
atgcattctta ccgctgcgta tgcgctagct acggaagctg ggtgccatat ccggttaagt
120

```

cagtatgctc ggaaagtccg ccagacgcag ttaagagtgg aatacctgcg ccttcggctg
 180
 gcgagcctgc ctgggtgtga tgctggcgcg gcagtaggaa ttgatcgctg actgcgttta
 240
 gatttcgaaa aaggactcac caaatcccag ggtcgacgag aagagtccat acccgctggc
 300
 gaagacgcca gcacgtataa cagacttatg aaagcgctgc gccaacgcca tgatgtcacc
 360
 aaatccggaa agcttgcccc
 379

<210> 48
 <211> 106
 <212> PRT
 <213> Homo sapiens

<400> 48
 Met His Leu Thr Ala Ala Tyr Ala Val Ala Thr Glu Ala Gly Cys His
 1 5 10 15
 Ile Arg Leu Ser Gln Tyr Ala Arg Lys Val Arg Gln Thr Gln Leu Arg
 20 25 30
 Val Glu Tyr Leu Arg Leu Arg Leu Ala Ser Leu Pro Gly Gly Asp Ala
 35 40 45
 Gly Ala Ala Val Gly Ile Asp Arg Arg Leu Arg Leu Asp Phe Glu Lys
 50 55 60
 Gly Leu Thr Lys Ser Gln Gly Arg Arg Glu Glu Phe Ile Pro Val Gly
 65 70 75 80
 Glu Asp Ala Ser Thr Tyr Asn Arg Leu Met Lys Ala Leu Arg Gln Arg
 85 90 95
 His Asp Val Ile Lys Ser Gly Lys Leu Ala
 100 105

<210> 49
 <211> 309
 <212> DNA
 <213> Homo sapiens

<400> 49
 tgatcatgat gctggcatgg actattctgg tcctctgtcc tctctcaect gctgaaggac
 60
 atccccetata tttttgtgtc tcctctgtga tcatacaatt ttcctctctc actgagtctc
 120
 ttgcattctc ttggaagcat gctgtactat gtcccatcct taaagaactc ccttgtctgt
 180
 cacattaccc tctgccagct ggctcatttt tctgtctccc ttacagggga aactettcaa
 240
 aaagtatatc ccacctcett ccattctcatg ttctcttgaa cctgcagtag tgggtgctcc
 300
 ctcccttttg
 309

<210> 50
 <211> 101
 <212> PRT

<213> Homo sapiens

<400> 50

```
Met Met Leu Ala Trp Thr Ile Leu Val Pro Val Pro Leu Ser Pro Ala
 1           5           10           15
Glu Gly His Pro Ser Asn Phe Cys Val Ser Phe Cys Ile Ile Lys Phe
          20           25           30
Ser Leu Ser Thr Glu Ser Leu Ala Ser Pro Trp Lys His Ala Val Leu
          35           40           45
Cys Pro Ile Leu Lys Glu Leu Pro Leu Ser Ala His Tyr Pro Leu Pro
          50           55           60
Ala Gly Ser Phe Phe Cys Ser Pro Leu Gln Gly Asn Ser Ser Lys Ser
65           70           75           80
Tyr Leu His Leu Leu Pro Ser His Val Leu Leu Asn Leu Gln Tyr Trp
          85           90           95
Val Leu Pro Pro Phe
          100
```

<210> 51

<211> 512

<212> DNA

<213> Homo sapiens

<400> 51

```
agatccttga agaattgcca cactgtcttc ctcctgctt ataatttcct tattccctag
60
gatgtgatcc ttgttcttgg ggcctcacat ggcagctgga tctctggcga ttgcactga
120
gttccagaca ccaggatgga aaagaaaaga agagggggca agaggaaacc ccagatgctc
180
cttaagagct actgcgtggc attcccactt gcactctcatt tgctcgatcg ctgtcactgt
240
gccctaaaga gctgcaagga cactggggaa atgagctctgt cttgtacttc atgtgcccct
300
caaaatcttc tgttgctgag ggagaagagg ccagccggta ttgaggaaca actagcactt
360
tctgcttcgg cgtcccaggg ggacgtgggt gtgttgaaac cacaccgggg gtgcggacct
420
ctgaggctgg gctggatggg acatcaggtg ggcctctctg ttcatttatg tgacctccca
480
tcaggctctc tggttggatc ctgctttcta ga
512
```

<210> 52

<211> 125

<212> PRT

<213> Homo sapiens

<400> 52

```
Met Glu Lys Lys Arg Arg Arg Gly Lys Arg Asn Pro Gln Met Leu Leu
 1           5           10           15
Lys Ser Tyr Cys Val Ala Phe Pro Leu Ala Ser His Leu Leu Asp Arg
          20           25           30
Cys His Cys Ala Leu Thr Ser Cys Lys Asp Thr Gly Glu Met Ser Leu
```

```

          35          40          45
Ser Cys Thr Ser Cys Ala Pro Gln Asn Leu Leu Leu Leu Arg Glu Lys
   50          55          60
Arg Pro Ala Gly Ile Glu Glu Gln Leu Ala Leu Ser Ala Ser Ala Ser
   65          70          75          80
Gln Gly Asp Val Gly Val Leu Asn Pro His Arg Gly Cys Gly Pro Leu
          85          90          95
Arg Leu Gly Trp Met Gly His Gln Val Gly Pro Leu Phe His Leu Cys
          100          105          110
Asp Leu Pro Ser Gly Leu Leu Val Gly Ser Cys Phe Leu
          115          120          125

```

<210> 53

<211> 474

<212> DNA

<213> Homo sapiens

<400> 53

```

accggtacac ctacgtcacc cgtaaaaacc gacgcaatac cgggatcgcc tcgtcctcaa
   60
aaaattcgat cccgtcgtgc gtcgtcacat tgagttcaag gaggcccgct aatggccaaa
   120
aagtccaaga ttgtcgccca gaagaaacgt gagaagctcg tagcccaata cgccgaaagg
   180
cgcgccgaac tcaaggccat catgaagtgc ccaactgcct cattggacga acgcatggag
   240
gcatcgcgta agctgtctcg cctgccgcgc gattcatccc ccgtgcggtt acgtaaccgt
   300
gaccaagtcg acgggcgtcc ccgcggctac gttggcaaagg ccggtgtgtc ccgtatccgt
   360
ttccgtgaga tggccaccgc cggcgaaactc cccggaatcg cgaagtcaag ctggtgaagc
   420
catggcagta ccgaagcgaa agaagtcccg ttgcaccacg cgtcataggc gggc
   474

```

<210> 54

<211> 101

<212> FRT

<213> Homo sapiens

<400> 54

```

Met Ala Lys Lys Ser Lys Ile Val Ala Gln Lys Lys Arg Glu Lys Leu
   1          5          10          15
Val Ala Gln Tyr Ala Glu Arg Arg Ala Glu Leu Lys Ala Ile Met Lys
          20          25          30
Cys Pro Thr Ala Ser Leu Asp Glu Arg Met Glu Ala Ser Arg Lys Leu
          35          40          45
Ser Arg Leu Pro Arg Asp Ser Ser Pro Val Arg Leu Arg Asn Arg Asp
          50          55          60
Gln Val Asp Gly Arg Pro Arg Gly Tyr Val Gly Lys Ala Gly Val Ser
   65          70          75          80
Arg Ile Arg Phe Arg Glu Met Ala His Arg Gly Glu Leu Pro Gly Ile
          85          90          95
Ala Lys Ser Ser Trp

```

100

<210> 55
 <211> 378
 <212> DNA
 <213> Homo sapiens

<400> 55
 ccattggccca ggacagccgg cataatcggtc acgactacgg tacaccgggtg ggcgccacagt
 60
 tcggcgccagc caagcccgca gcgtgctgcc aggcgcaagc gacaaacacc ggcgcgtggg
 120
 tgggtgttcga ccatgtgcgt tgcacccacg acacctttct gatcgacgtc tttctcaacc
 180
 agcccgatgc caccgcgcag caggtcaatg ccgacaaccc gcactacgtc gggcgtttca
 240
 gccgcacgtc catgggcctg gtggatgaca agggcggttg cattaccacg ggcgtatcgc
 300
 gcgcgttgaa tgccggcgcg agcaccaagg cgctgaacct gggaccgagt gacgcggcgc
 360
 agttatcggg gaggcgta
 378

<210> 56
 <211> 125
 <212> PRT
 <213> Homo sapiens

<400> 56
 Met Ala Gln Asp Ser Arg His Ile Gly Tyr Asp Tyr Gly Thr Pro Val
 1 5 10 15
 Ala Pro Gln Phe Gly Ala Ala Lys Pro Ala Ala Cys Cys Gln Ala Gln
 20 25 30
 Ala Thr Asn Thr Gly Pro Trp Val Val Phe Asp His Val Arg Cys Thr
 35 40 45
 His Asp Thr Phe Leu Ile Asp Val Phe Leu Asn Gln Pro Asp Ala Thr
 50 55 60
 Ala Gln Gln Val Asn Ala Asp Asn Pro His Tyr Val Gly Arg Phe Ser
 65 70 75 80
 Arg Ile Gly Met Gly Leu Val Asp Asp Lys Gly Arg Cys Ile Thr Gln
 85 90 95
 Gly Val Ser Arg Ala Leu Asn Ala Ala Arg Ser Thr Lys Ala Leu Asn
 100 105 110
 Leu Gly Pro Ser Asp Ala Ala Gln Leu Ser Val Arg Arg
 115 120 125

<210> 57
 <211> 388
 <212> DNA
 <213> Homo sapiens

<400> 57
 agaccacccc gacacagatc aggagtcgtc atgtccagaa agaagaaggt cggcatcctc
 60

accgcaggcg gtgattgccc cgggctcaac gccgctatcc gccgatttgg caaggctgcc
 120
 atccgccagc acgacatgga gctcatcggt attcaggacg gctttcttgg attggcggga
 180
 aaccgcacca tctcccttgg cccgcgtgcc ctctcaggca tcttgacggt cggcggggac
 240
 atcctgggaa ctaccgctga caaggtcaat cacatgatta tcgacggcga ggaacgggat
 300
 atgggtcccca ccaccgtoga gaattacgag aagctggggc ttgacgcttt ggtgactttg
 360
 ggtggcgggt gcaccgcca gaacgcgt
 388

<210> 58
 <211> 129
 <212> PRT
 <213> Homo sapiens

<400> 58
 Arg Pro Thr Arg His Arg Ser Gly Val Val Met Ser Arg Lys Lys Lys
 1 5 10 15
 Val Gly Ile Leu Thr Ala Gly Gly Asp Cys Pro Gly Leu Asn Ala Ala
 20 25 30
 Ile Arg Gly Phe Gly Lys Ala Ala Ile Arg Gln His Asp Met Glu Leu
 35 40 45
 Ile Gly Ile Gln Asp Gly Phe Leu Gly Leu Ala Gly Asn Arg Thr Ile
 50 55 60
 Ser Leu Gly Pro Arg Ala Leu Ser Gly Ile Leu Thr Val Gly Gly Thr
 65 70 75 80
 Ile Leu Gly Thr Ser Arg Asp Lys Val Asn His Met Ile Ile Asp Gly
 85 90 95
 Glu Glu Arg Asp Met Val Pro Thr Thr Val Glu Asn Tyr Glu Lys Leu
 100 105 110
 Gly Leu Asp Ala Leu Val Thr Leu Gly Gly Gly Thr Ala Lys Asn
 115 120 125
 Ala

<210> 59
 <211> 417
 <212> DNA
 <213> Homo sapiens

<400> 59
 ggtaccatcg gagctcgaca agaaatggtt ggggaagtc gtggcttctg ctccaccag
 60
 tgccctcatg ggtcagccca cctgaatatc ttcatgcctg tgcattttct ctgatgttca
 120
 cgtgtgcctt gtgtttttac gcatctgtga tcgtgcaccc accgctctca gagaggagcc
 180
 cgtttgggaa tcgggagaat gtgcgctggc ggaagacgct cacacactgg aagcaaacct
 240
 cagaccgcgt ggacaagacc aaggatgaaa tggaacacga ggccttggtg gaagggaacc
 300

tggcaaccga ggcaagccta gtggttctgg acacactgga gatcatcgtg cagacgggtga
 360
 tgctttcaga agccccgggag agcgtcttgg gggcagtgct gaaggttgtg ctgtaca
 417

<210> 60
 <211> 101
 <212> PRT
 <213> Homo sapiens

<400> 60
 Met Phe Thr Cys Ala Leu Cys Phe Tyr Ala Ser Val Ile Val His Pro
 1 5 10 15
 Arg Val Ser Glu Arg Ser Pro Phe Gly Asn Pro Glu Asn Val Arg Trp
 20 25 30
 Arg Lys Ser Val Thr His Trp Lys Gln Thr Ser Asp Arg Val Asp Lys
 35 40 45
 Thr Lys Asp Glu Met Glu His Glu Ala Leu Val Glu Gly Asn Leu Ala
 50 55 60
 Thr Glu Ala Ser Leu Val Val Leu Asp Thr Leu Glu Ile Ile Val Gln
 65 70 75 80
 Thr Val Met Leu Ser Glu Ala Arg Glu Ser Val Leu Gly Ala Val Leu
 85 90 95
 Lys Val Val Leu Tyr
 100

<210> 61
 <211> 304
 <212> DNA
 <213> Homo sapiens

<400> 61
 agatcttcac agccttagac ttttttcatg ggtgccttac agttttggag gtccttatcc
 60
 gcacacatat ttgcaggctt ggagagagtg tgtgggggca tgtactttcg gtgggtcaag
 120
 tatgaagaag caggccttat aaacacatat tctgacctta acctgtactt cagaagagga
 180
 ccgctgactc accaaggagg cctgaaggac aaggcagcat ctctgtcttc acatgagtc
 240
 tccccctagac cggggcccatg gccaggccctg accacagagc tcccattgcc tttcctgcac
 300
 gcgt
 304

<210> 62
 <211> 92
 <212> PRT
 <213> Homo sapiens

<400> 62
 Met Gly Ala Leu Gln Phe Trp Arg Ser Leu Ser Ala His Ile Phe Ala
 1 5 10 15

```

Gly Leu Glu Arg Val Cys Gly Gly Met Tyr Phe Arg Trp Val Lys Tyr
20          25          30
Glu Glu Ala Gly Leu Ile Asn Thr Tyr Ser Asp Leu Asn Leu Tyr Phe
          35          40          45
Arg Arg Gly Pro Leu Thr His Gln Gly Gly Leu Lys Asp Lys Ala Ala
          50          55          60
Ser Leu Ser Ser His Glu Ser Ser Pro Arg Pro Gly Pro Trp Pro Gly
65          70          75          80
Leu Thr Thr Glu Leu Pro Leu Pro Phe Leu His Ala
          85          90

```

<210> 63

<211> 577

<212> DNA

<213> Homo sapiens

<400> 63

```

cgcgctcaagg ggggtctacac cgggacgatt aacgcctcgg tgggagtagt catcacccgg
60
ctgacgggtgc tagctgggtg gctcacecta gccggggcta tcagtgtcgg ggaactcgtc
120
accgtgggtcg ggctggccca aaccctoggc cctcgcgtgc gagcactggg cgtcgacacc
180
ggcagcatgt tggccaccgc ccacgcctcc ggggaccgat tctgtgagtt gcgtgatagc
240
ccggcagcct ggcagatcca ccccgacgac ggtgcccgca ccacaccggg tgatggcccg
300
gtggagttgc acatcccggt cagggatttc cagcttgacg tcgccggcgg caccatgtg
360
ggtatcatgg cgcctcaatc ggtctgtgac gccttggccg aggcgataga ccacggctcc
420
gagacggtct tgaatggggg tcccgcagtg cgcctcaacc ctgcccacgg gcgtcgtctg
480
gtgctgtgtg ctccccgctc ccccgaaactg ttgcagcata ctgcccgctg gaacatcggt
540
cttgacagcc agacgactgt cgccaggctg aatgcat
577

```

<210> 64

<211> 192

<212> PRT

<213> Homo sapiens

<400> 64

```

Arg Val Lys Gly Val Tyr Thr Gly Thr Ile Asn Ala Ser Val Gly Val
1          5          10          15
Phe Ile Thr Ala Leu Thr Val Leu Ala Gly Trp Leu Thr Leu Ala Gly
          20          25          30
Arg Ile Ser Val Gly Glu Leu Val Thr Val Val Gly Leu Ala Gln Thr
          35          40          45
Leu Gly Pro Pro Leu Arg Ala Leu Gly Val Asp Thr Ala Thr Met Leu
          50          55          60
Ala Thr Ala His Ala Ser Gly Asp Arg Phe Cys Glu Leu Arg Asp Ser
65          70          75          80

```



```

Pro Ala Ala Trp Gln Ile His Pro Asp Asp Gly Ala Arg Thr Thr Pro
      85          90          95
Gly Asp Gly Pro Val Glu Leu His Ile Pro Val Arg Asp Phe Gln Leu
      100        105        110
Asp Val Ala Gly Gly Thr His Val Gly Ile Met Ala Pro Gln Ser Val
      115        120        125
Cys Asp Ala Leu Ala Glu Ala Ile Asp His Gly Ser Glu Thr Val Leu
      130        135        140
Asn Gly Val Pro Ala Ser Arg Leu Asn Pro Ala Gln Arg Arg Arg Leu
      145        150        155        160
Val Leu Val Ala Pro Arg Ser Pro Glu Leu Phe Asp Asp Thr Ala Arg
      165        170        175
Ala Asn Ile Val Leu Asp Ser Gln Thr Thr Val Ala Arg Leu Asn Ala
      180        185        190

```

<210> 65

<211> 339

<212> DNA

<213> Homo sapiens

<400> 65

```

gtcgaccgcg ccttgggcatc gctcgaaggc gccagcctgg accaggtagc ggaagaagtc
60
aagaaggccg ctttcaagat caccgcgcgc ggcgaactag tgggcaccat ggcctccgag
120
cgcccttgcg tacccttcgg catcatcgac ctttcgcttg ccctactgc cgaattggga
180
gattcggggg ccacatcct tgagcatatg ggattggacc aagtaggcac gcacggcaca
240
actgctgctt tggctctgct taacgacgcc gtaaagaaa gcgcatgat ggcctgcccc
300
cgcgctggcg gtttgtctgg ctcttcctac ccgggctcc
339

```

<210> 66

<211> 113

<212> PRT

<213> Homo sapiens

<400> 66

```

Val Asp Arg Ala Leu Gly Ser Leu Glu Gly Ala Ser Leu Asp Gln Val
1      5      10      15
Ala Glu Glu Val Lys Lys Ala Ala Phe Lys Ile Thr Arg Ala Gly Gln
20     25     30
Leu Val Gly Thr Met Ala Ser Glu Arg Leu Gly Val Pro Phe Gly Ile
35     40     45
Ile Asp Leu Ser Leu Ala Pro Thr Ala Glu Leu Gly Asp Ser Gly Ala
50     55     60
His Ile Leu Glu His Met Gly Leu Asp Gln Val Gly Thr His Gly Thr
65     70     75     80
Thr Ala Ala Leu Ala Leu Leu Asn Asp Ala Val Lys Lys Gly Gly Met
85     90     95
Met Ala Cys Pro Arg Val Gly Gly Leu Ser Gly Ser Phe Ile Pro Gly
100    105    110

```

Ser

<210> 67
 <211> 446
 <212> DNA
 <213> Homo sapiens

<400> 67
 tgatcataaa ccaacggtca cagaggggat gtggcacacc tacctgcgcg tcgcagatgc
 60
 cgcacaggca cgggtcaggg gcgttcgcgg cgccagctgg cacaacttcg cgaccggtga
 120
 caaggggtcc ttgcacgcca acgagcttgc cgttaactct gatactgaca ccgtcatcca
 180
 gggagtcggg ccgcacctag ccctcctcga ttcagcgtgg ggaagccaga tccacgtgga
 240
 gacaacaggg tggtccagtg ccgtgggtctg gaatccacgc tctcgtcga cacatgcgga
 300
 taaccgaca gccacggcat ggcgcgattt cgtatgcgtc gagaccgggg cctgcaagga
 360
 caatgcggtc attgttgccc cacacagcga cctcaccatg tccacacgga ttagegtcga
 420
 aacgttgtga tcgctgcatg gatatt
 446

<210> 68
 <211> 133
 <212> PRT
 <213> Homo sapiens

<400> 68
 Met Trp His Thr Tyr Leu Arg Val Ala Asp Ala Ala Gln Ala Arg Val
 1 5 10 15
 Arg Gly Val Arg Gly Ala Ser Trp His Asn Phe Ala Thr Gly Asp Lys
 20 25 30
 Gly Ser Phe Asp Ala Asn Glu Leu Ala Val Thr Pro Asp Thr Asp Thr
 35 40 45
 Val Ile Gln Gly Val Gly Pro Ala Leu Ala Leu Leu Asp Ser Ala Trp
 50 55 60
 Gly Arg Gln Ile His Val Glu Thr Thr Gly Cys Pro Ser Ala Val Val
 65 70 75 80
 Trp Asn Pro Arg Ser Ser Ser Thr His Ala Asp Asn Pro Thr Ala Gln
 85 90 95
 Ala Trp Arg Asp Phe Val Cys Val Glu Thr Gly Ala Cys Lys Asp Asn
 100 105 110
 Ala Val Ile Val Ala Pro His Ser Asp Leu Thr Met Ser Thr Arg Ile
 115 120 125
 Ser Val Glu Thr Leu
 130

<210> 69
 <211> 552
 <212> DNA

<213> Homo sapiens

<400> 69

nnagggttaa ggagaaaagc aaggaccttg caaagagagc ctctgtgccg gagaggctgg
60
ccctcaagga ggagccaaaa gaagacccca gtggagcagc tgtgcccgag atgccaaaaa
120
agtcctccaa gattgccagc ttcacccca aaggggggaa gctcaacagt gccagaagga
180
agcncatgg ccccttcctt cagtggaaata ccaaaaccag gaatgaaaag catgcccggg
240
aaatccccaa gtgccccagc gccttccaag gaagggggagc ggagccggag tgggaagctg
300
agctcaggac tccccagca gaagccccag ctggacggca gacactccag ttcctcttcc
360
agcctggcgt cctcagaagg aaaaggccca ggagggacca cctgaacca cagcatcagc
420
agccagactg tcagtgggtc tgcggggacc acccagacca caggaagcaa tnnaccgtca
480
gtgttcagct acctcagccc cagcagcaat acaaccatcc caaactgcc acggttgccac
540
ctttctgta ca
552

<210> 70

<211> 184

<212> PRT

<213> Homo sapiens

<400> 70

Xaa	Arg	Val	Arg	Arg	Lys	Ala	Arg	Thr	Leu	Gln	Arg	Glu	Pro	Leu	Cys
1				5					10					15	
Arg	Arg	Gly	Trp	Pro	Ser	Arg	Arg	Ser	Gln	Lys	Lys	Thr	Pro	Val	Glu
			20					25					30		
Gln	Leu	Cys	Pro	Arg	Cys	Gln	Lys	Ser	Pro	Pro	Arg	Leu	Pro	Ala	Ser
		35					40					45			
Ser	Pro	Lys	Gly	Gly	Ser	Ser	Thr	Val	Pro	Arg	Arg	Ser	Xaa	Met	Ala
	50				55					60					
Pro	Ser	Leu	Ser	Gly	Ile	Pro	Lys	Pro	Gly	Met	Lys	Ser	Met	Pro	Gly
65				70						75				80	
Lys	Ser	Pro	Ser	Ala	Pro	Ala	Pro	Ser	Lys	Glu	Gly	Glu	Arg	Ser	Arg
			85						90					95	
Ser	Gly	Lys	Leu	Ser	Ser	Gly	Leu	Pro	Gln	Gln	Lys	Pro	Gln	Leu	Asp
			100					105					110		
Gly	Arg	His	Ser	Ser	Ser	Ser	Ser	Ser	Leu	Ala	Ser	Ser	Glu	Gly	Lys
		115					120					125			
Gly	Pro	Gly	Gly	Thr	Thr	Leu	Asn	His	Ser	Ile	Ser	Ser	Gln	Thr	Val
		130				135					140				
Ser	Gly	Ser	Val	Gly	Thr	Gln	Thr	Thr	Gly	Ser	Asn	Xaa	Pro	Ser	
145				150					155				160		
Val	Phe	Ser	Tyr	Leu	Ser	Pro	Ser	Ser	Asn	Thr	Thr	Ile	Pro	Thr	Leu
			165					170					175		
Pro	Arg	Leu	His	Leu	Ser	Cys	Thr								
			180												

<210> 71
 <211> 316
 <212> DNA
 <213> Homo sapiens

<400> 71
 cgcgttgaaa tggcgttcga acttaaacgt ttacatattg actccgtgcc attaacatt
 60
 ttgaatcctg ttaaaggga ctcatttgaa agcaacgaag ctttacgtcc tttaaatata
 120
 ttacgtacct tcgcccgtatt ccgtttcacc ttgccaaacg cattgatacg aactgcagg
 180
 ggcccgaag taaatctacg agacttacaa gcttatgctc taaaagggtg cctaaacggt
 240
 atcatgggtg ttggctactt aactactggc ggtcgttcac ctcaagacga tctccaaatg
 300
 attcaagact tggagt
 316

<210> 72
 <211> 105
 <212> PRT
 <213> Homo sapiens

<400> 72
 Arg Val Glu Met Ala Phe Glu Leu Lys Arg Leu His Ile Asp Ser Val
 1 5 10 15
 Pro Leu Asn Ile Leu Asn Pro Val Lys Gly Thr Pro Phe Glu Ser Asn
 20 25 30
 Glu Ala Leu Arg Pro Leu Asn Ile Leu Arg Thr Phe Ala Val Phe Arg
 35 40 45
 Phe Ile Leu Pro Asn Ala Leu Ile Arg Thr Ala Gly Gly Arg Glu Val
 50 55 60
 Asn Leu Arg Asp Leu Gln Ala Tyr Ala Leu Lys Gly Gly Leu Asn Gly
 65 70 75 80
 Ile Met Val Gly Gly Tyr Leu Thr Thr Gly Gly Arg Ser Pro Gln Asp
 85 90 95
 Asp Leu Gln Met Ile Gln Asp Leu Glu
 100 105

<210> 73
 <211> 384
 <212> DNA
 <213> Homo sapiens

<400> 73
 nntaccggca agatcctggc cgaagggtgac gtcagggttt ctgaggctat cgactttgct
 60
 gcttgggtatg tcgaccgagc cgaggagctc gagggcgctg acgggtgccca gtttgtgccg
 120
 ccacgagtga ccgtcgtcac cccgccgtgg aacttcgccc tgcctattac cgcgggatac
 180

acccttgccg ctctggccgc cggatcgta gtactactca agcccgctcc acaggcccg
 240
 cactgtgctg ccgtcatctc tgaatgcctg tgggaggctg ggatcccgcg ggaagttctg
 300
 cagctcgctg atgttgagga aaatgaggct ggtaaacacc tggtagacca ccccgaggct
 360
 gatcgggtca tcctcacggg aggt
 384

<210> 74

<211> 128

<212> PRT

<213> Homo sapiens

<400> 74

Xaa	Thr	Gly	Lys	Ile	Leu	Ala	Glu	Gly	Asp	Val	Glu	Val	Ser	Glu	Ala
1			5					10						15	
Ile	Asp	Phe	Ala	Ala	Trp	Tyr	Val	Asp	Arg	Ala	Glu	Glu	Leu	Glu	Gly
			20					25						30	
Val	Asp	Gly	Ala	Gln	Phe	Val	Pro	Arg	Val	Thr	Val	Val	Thr	Pro	
		35					40				45				
Pro	Trp	Asn	Phe	Ala	Leu	Ser	Ile	Thr	Ala	Gly	Ser	Thr	Leu	Ala	Ala
		50				55					60				
Leu	Ala	Ala	Gly	Ser	Ser	Val	Leu	Leu	Lys	Pro	Ala	Pro	Gln	Ala	Arg
		65			70				75					80	
His	Cys	Ala	Ala	Val	Ile	Ser	Glu	Cys	Leu	Trp	Glu	Ala	Gly	Ile	Pro
		85						90						95	
Arg	Asp	Val	Leu	Gln	Leu	Val	Asp	Val	Glu	Glu	Asn	Glu	Ala	Gly	Lys
		100						105						110	
His	Leu	Val	Ser	His	Pro	Glu	Val	Asp	Arg	Val	Ile	Leu	Thr	Gly	Gly
		115					120							125	

<210> 75

<211> 405

<212> DNA

<213> Homo sapiens

<400> 75

gaattcgtct cggaatacac gctggaaaaat tcggccgaga tgtccggggt gcgctcanac
 60
 cgcattgagg cgctggccga gctctatgcc gatcccaaga ccagggtggt gagcttctgg
 120
 accatgggct tcaaccagca caccgcggc gtctgggtgca acaatctcgt ctacaacatc
 180
 cacctgctga ccggaaaaat ctgcacgccc ggcaacagcc cgttctcgtg gaccgggagc
 240
 ccattcgctt cgggcacggc gcgcgaggtc ggtacctctt cgcctcgctt gcccgccgac
 300
 atgggtgtgta ccagcaaggc gcaccgcgac atcgccgaga agatctggca gctgccggaa
 360
 ggaccagtcc ccgacaagcc cggctaccac gccgtgctgc agagc
 405

<210> 76

<211> 135
 <212> PRT
 <213> Homo sapiens

<400> 76
 Glu Phe Val Ser Glu Tyr Thr Leu Glu Asn Ser Ala Glu Met Ser Gly
 1 5 10 15
 Val Arg Ser Xaa Arg Ile Glu Ala Leu Ala Glu Leu Tyr Ala Asp Pro
 20 25 30
 Lys Thr Arg Val Val Ser Phe Trp Thr Met Gly Phe Asn Gln His Thr
 35 40 45
 Arg Gly Val Trp Cys Asn Asn Leu Val Tyr Asn Ile His Leu Leu Thr
 50 55 60
 Gly Lys Ile Ser Thr Pro Gly Asn Ser Pro Phe Ser Leu Thr Gly Gln
 65 70 75 80
 Pro Ser Ala Cys Gly Thr Ala Arg Glu Val Gly Thr Phe Ser His Arg
 85 90 95
 Leu Pro Ala Asp Met Val Val Thr Ser Lys Ala His Arg Asp Ile Ala
 100 105 110
 Glu Lys Ile Trp Gln Leu Pro Glu Gly Pro Val Pro Asp Lys Pro Gly
 115 120 125
 Tyr His Ala Val Leu Gln Ser
 130 135

<210> 77
 <211> 5816
 <212> DNA
 <213> Homo sapiens

<400> 77
 gagcggcgcc ctgctctggc cgttgctccc gctcctgctc ctgctgctgt cggcgcggga
 60
 cggcgctgcgc gccgcgcagc ctcaggcccc ggggttacttg attgcagctc cctctgtttt
 120
 tcgcgcggcg gtggaggaag tcatcagcgt gaccatcttt aactctccaa gggaagtcac
 180
 ggtccaggct cagctggttg cccagggtga gccggtggtg cagagccagg gagccatcct
 240
 ggataaaggg acaatcaaac tcaagggtgcc caggggcctc cggggccaag cgcttctgaa
 300
 agtgtgtggc cgcggtggc aggcggagga ggggcccctc ttccacaacc agacctcggt
 360
 gaccgtggac ggccggggcg cttctgtatt catccagacg gacaagcctg tgtacagacc
 420
 ccagcaccga gtgctcataa gcatcttcac cgtctctcca aatctgaggc ctgtcaacga
 480
 gaagctggaa gcctacatcc tggacccccg aggtctctcg atgatatagt ggagacactt
 540
 aaagccgttc tgctgcggca tcaccaacat gagcttcccc ttgtccgacc agcctgtggt
 600
 gggagaatgg ttcatttttg ttgaaatgca aggccacgag tacaacaagt cttttgaagt
 660
 tcagaagtat gtgttgccca agtttgagct tctgattgac ccgccccggt atatccaaga
 720

cctggacgco tgtgagacag gcactgtgco ggccaggat acctttggga aacctgtggc
 780
 tggtagccta atgatcaaca tgactgttaa tgggttaggg tactacagcc acgaggtaggg
 840
 acgccctgtc ctcagaacaa ccaagatcct cggctcccgg gacttcgaca tctcgtgag
 900
 ggacatgatc ccagcggacg tccctgagca cttccggggc agggtagca tctgggcat
 960
 ggtgaccagt gtggacggga gccagcaggc cgcgttcgat gactccaccc cgtgacagag
 1020
 gcagctgggt gacatccggc actccaagga cagcaggaa cagttcaagc cgggctgggc
 1080
 ctacgtgggg aagggtggagc tatcctaccc cgatggcagc ccagctgagg gggtagcggc
 1140
 ccagattaag gcagagctga caccaaagga taacatctac accagtgaag ttgtgtcca
 1200
 gcgtggacta gtggggtttg aaatcccctc catccccacg tcagcccagc acgtgtggct
 1260
 ggagaccaag gtgatggcac tgaacgggaa gcccgtaggg gctcagtacc tgcccagcta
 1320
 cctctccctc ggcagctggc actccccag ccagtgtac ctgcagctgc agccacccctc
 1380
 ccaccactcg cagggtgggg aagaagccta tttttctgtg aagtcacat gtccctgcaa
 1440
 ctttaccctg tactacgagg tggctgcacg gggcaatatt gtgctatcg gccagcagcc
 1500
 tgcccacacc acccagcagc gaagcaagcg ggcggcccct gccctggaga aaccgattcg
 1560
 tttaacacac ctttctgaga cagagccccc accagcccca gaagctgagg tcgacgtgtg
 1620
 tgtgacctct cttcatctgg ccgtgacccc cagcatgggc ccccttgggc gcctgtcgtg
 1680
 cttctacgtc agggagaatg gagaaggggt ogccgacagc cttcagtttg cagtgcagac
 1740
 cttctcgaa aaccaggttt cagtgcagta tcagcaaat gagaccacac ctggggaggt
 1800
 tgtcagctcg cggatcaggg ctgcaagggg cagctgtgtg tgcgtcgccg cagttgataa
 1860
 gagtgtctac ctgctcaggt ctgggttcgc gcgactcct gcccagggtt tccagggaat
 1920
 ggaagattat gatgtttctg attccttttg cgtgtccagg gaggatgggc ctttttgggtg
 1980
 ggctgggctg acggcacaac gacgcggcg ctcctctgtc ttcccgtagc cttggggcat
 2040
 caccaaggac tctgggtttg ctttcaccga aacgggagct gtggtgatga ccgaccagat
 2100
 gagcctgaac caccggcagg acggtggcct ctacaccgat gaggtgtgct ccgctttcca
 2160
 gcccacaca gggagcctgg tggcagtgcc tcttccagg cccccccca gaacagagaa
 2220
 gagaaaaagg actttcttcc cgaaaacatg gatttggcat tgtctcaaca tcagtgaccc
 2280
 atctgggtgag gggacactca gtgtgaaggt cccggagctc atcaccagct ggggtgggtga
 2340

ggcgcgtggcc ctgtccacct ctacaggcctt aggcacgcgc gagccctccc tgctgaagac
2400
cttcaagccc ttcttcgttg acttcacgct ccccgctctc atcatccgtg gggagcagggt
2460
caagatcccc ctacgtgtct acaactacat gggcacctgc gctgagggtg acatgaagct
2520
ctcgggtccc aagggcatcc agttttgttg gcatcctggc aaacgccatg tgaccaagaa
2580
gatgtgtgtg gcccccgggg aggcctgagcc catctgggtc gttctgtcct tcacgcaccc
2640
gggactcaac aacatcacgg ccaaagccct tgcttacgga gacacaaatt gctgccggga
2700
tgggagggtcc agcaaacacc ctgaggagaa tcacgccgac aggagggtcc ccacgcgggt
2760
ggatcacgtc aggcgcagtg tgatggttga ggccgaagga gtcccccggt cgtaacaccta
2820
cagcgcattc ttctgtccca gtgagagagt ccacatctcc accccaaca agtatgagtt
2880
ccagtatgtg cagcggccac tgcgcctcac ccgctttgat gtggctgtgc gagctcacia
2940
tgatgccctg gtggccttgt cttctggggc ccaggacaca gcaggcatga tcgagatcgt
3000
cctggggggg catcagaaca ccaggctcatg gatctccacc agcaagatgg gagagcccg
3060
ggccaagtga cacacggcca agatcctctc ctgggatgaa ttcagaacat tctggatcag
3120
ctggcgtggt ggccttatcc aggttggcca tggctccagc ccacccaatg agtctgtcat
3180
tgtggccttg accctcccga gggcaccaga ggtccagttc attggctttt ccacgcggtg
3240
gggctccatg ggtgaattcc gaatctggag gaagatggag gtggacgaga gctacagcga
3300
ggccttcacc tgggggtccc acacggcgcc atccctgggt ctgagcgagc caccgcctcc
3360
atcatcgggg acgtcatggg gcccaaccctg aaccacctca acaacctcct gcggtgcgg
3420
tttggctgtg gagagcagaa catgatccac ttgacacca acgtctttgt ctgaagtat
3480
cttcagaaaa ccacagcagct cagccctgag gtggagagag agaccacaga ctacctagta
3540
caaggctacc agcgcacagct gacctacaag cgcaggatg gctcctacag cgcgtttggg
3600
gagcgggacg catcggggag catgtggctc acagcctttg tcctgaagtc cttgcacacg
3660
gtcgcagctg ttatcttcgt ggacccccgg gagctggctg ccgccaagag ctggatcctc
3720
cagcagcagc agggccgatg ctccctcctg gccctgggca gggctctgaa caaggacatc
3780
cagggtggga tccacggcat tgtcccgtg acagcctacg tgggtgttgc tctcctggaa
3840
acaggcacag cctcagagga ggagagaggc tccactgaca aagcgaggca cttcctggag
3900
tctgctgcgc ccctggccat ggacccctat agctgtgccc tgactacctc cgcgtgacc
3960

ctgctccgca gccccgcagc ccctgaggca ctgcgcaagc tccgtagcct ggccatcatg
 4020
 cgagatgggg tcacccactg gaggctgtca aattctctgg acgtggacaa gggcacattc
 4080
 ttgagcttca gtgacagggt ctctcagtcg gtgggtctcg ccgaggtgga aatgacagcc
 4140
 tacgcccctc tgacctacac tctgctgggt gacgtggctg ccgccccgcc tgtgggtgaag
 4200
 tggctgtccc agcagcgaaa tgcacttggg ggcttctcct ccactcagga cacctgcgtg
 4260
 gctctgcagg ccttggtgta atatgccatc ttgtctctatg ctggaggcat caacctcact
 4320
 gtctctctgg cctccacca cctggactac caggaaacct tcgagctgca caggaccaac
 4380
 cagaagggtc tgcagacagc agcgatcccc agcctcccca cggggctggt tgtgagtgc
 4440
 aagggggcag gctgctgctt gatgcagatt gatgtcacct acaatgtgcc tgacccgggtg
 4500
 gccaaagccag ctttccagct gctcgtaagc ctccaggagc ctgaggccca gggacgcccc
 4560
 cccccctatg ctgcctccgc agctgagggt tcccaggagg actggcccc agctgacgat
 4620
 gatgaccagg cggccgatca gcatcaccag gaatacaagg tgatgctgga ggtgtgcacc
 4680
 aggtggctgc atgcagggtc ttccaatatg gctgtcctgg aggtgcccc gctgtcaggc
 4740
 ttccgggcag acatcgagag cctggagcag ctgctccttg acaagcacat ggggatgaag
 4800
 aggtatgaag tggctggagc ccgagtgctc ttctactttg atgagatccc cagccggtgc
 4860
 ctgacgtgcg tgcggttccg tgctctccgg gagtgcgtgg tgggcaggac gtcggcgctg
 4920
 ccagttctcc tgtacgacta ctacgaaccc gccttcgagg ccactcgctt ctacaacgtc
 4980
 agcacgcaca gccccactgc ccgggaactg tgcgcgggac ccgcgtgcaa cgaagtggag
 5040
 cgcgccccct cccggggccc gggctgggtc cccggcgagt cgggccccgc cgtggcccc
 5100
 gaggaggggg cgcgatcgc gcgatgcggc tgcgaccacg actgcggcgc ccagggggac
 5160
 ccggtgtgcg gctccgacgg ggtgggtctac gccagcgctt gccgcctgcg ggaggccggc
 5220
 tgcgccagg ccgcgcccc ggagcccgcg cctcccagct gctgcgccct cgagcagcgg
 5280
 ctgcggcct cgtcgtctc cactacggg gatgacctgg cttctgtggc cccggggcct
 5340
 ttacagcagg acgtgaagct gaatggagcc ggccctgagg tggaggactc agacctgag
 5400
 cctgaagggg aggcggagga cagggtcaca gccgggcctc ggctctctgt gagcagcggg
 5460
 aacctgaaaa cgacacccca gagcgccagc ccgttcacac gatggggcca gactccggcc
 5520
 cctcagagac atagtggcgg ggtgggtggg gccccacagc cagggtctct gagcctgtgc
 5580

ttcgtctaca gccacgcctt tcagagtggg ggggaggagg gtttatggat gtcaaacacc
 5640
 tgcaccttga gataatccta caaccacatg cagttgtggg accgcagttt ggctctgggg
 5700
 accattcata cccacacacc cagcttgtgc ctgtggttaa catctcagaa aactctggta
 5760
 aatgatcact ccaggatatt gacacgaata cagcttactg atcttactca catggt
 5816

<210> 78

<211> 799

<212> PRT

<213> Homo sapiens

<400> 78

Ile	Pro	Asn	Leu	Glu	Glu	Asp	Gly	Gly	Gly	Arg	Glu	Leu	Gln	Arg	Gly
1				5					10					15	
Leu	His	Leu	Gly	Val	Pro	His	Gly	Ala	Ile	Pro	Gly	Ser	Glu	Arg	Ala
			20					25					30		
Thr	Ala	Ser	Ile	Ile	Gly	Asp	Val	Met	Gly	Pro	Thr	Leu	Asn	His	Leu
		35					40					45			
Asn	Asn	Leu	Leu	Arg	Leu	Pro	Phe	Gly	Cys	Gly	Glu	Gln	Asn	Met	Ile
		50				55				60					
His	Phe	Ala	Pro	Asn	Val	Phe	Val	Leu	Lys	Tyr	Leu	Gln	Lys	Thr	Gln
65				70					75				80		
Gln	Leu	Ser	Pro	Glu	Val	Glu	Arg	Glu	Thr	Thr	Asp	Tyr	Leu	Val	Gln
			85						90				95		
Gly	Tyr	Gln	Arg	Gln	Leu	Thr	Tyr	Lys	Arg	Gln	Asp	Gly	Ser	Tyr	Ser
		100						105					110		
Ala	Phe	Gly	Glu	Arg	Asp	Ala	Ser	Gly	Ser	Met	Trp	Leu	Thr	Ala	Phe
		115					120					125			
Val	Leu	Lys	Ser	Phe	Ala	Gln	Ala	Arg	Ser	Phe	Ile	Phe	Val	Asp	Pro
		130				135					140				
Arg	Glu	Leu	Ala	Ala	Ala	Lys	Ser	Trp	Ile	Ile	Gln	Gln	Gln	Ala	
145				150						155			160		
Asp	Gly	Ser	Phe	Leu	Ala	Val	Gly	Arg	Val	Leu	Asn	Lys	Asp	Ile	Gln
			165					170					175		
Gly	Gly	Ile	His	Gly	Ile	Val	Pro	Leu	Thr	Ala	Tyr	Val	Val	Val	Ala
		180						185					190		
Leu	Leu	Glu	Thr	Gly	Thr	Ala	Ser	Glu	Glu	Arg	Gly	Ser	Thr	Asp	
		195					200				205				
Lys	Arg	His	Phe	Leu	Glu	Ser	Ala	Ala	Pro	Leu	Ala	Met	Asp	Pro	
	210				215					220					
Tyr	Ser	Cys	Ala	Leu	Thr	Thr	Tyr	Ala	Leu	Thr	Leu	Leu	Arg	Ser	Pro
225				230						235				240	
Ala	Ala	Pro	Glu	Ala	Leu	Arg	Lys	Leu	Arg	Ser	Leu	Ala	Ile	Met	Arg
			245					250					255		
Asp	Gly	Val	Thr	His	Trp	Ser	Leu	Ser	Asn	Ser	Trp	Asp	Val	Asp	Lys
			260					265					270		
Gly	Thr	Phe	Leu	Ser	Phe	Ser	Asp	Arg	Val	Ser	Gln	Ser	Val	Val	Ser
		275					280						285		
Ala	Glu	Val	Glu	Met	Thr	Ala	Tyr	Ala	Leu	Leu	Thr	Tyr	Thr	Leu	Leu
	290				295					300					
Gly	Asp	Val	Ala	Ala	Ala	Leu	Pro	Val	Val	Lys	Trp	Leu	Ser	Gln	Gln

```

305          310          315          320
Arg Asn Ala Leu Gly Gly Phe Ser Ser Thr Gln Asp Thr Cys Val Ala
          325          330          335
Leu Gln Ala Leu Ala Glu Tyr Ala Ile Leu Ser Tyr Ala Gly Gly Ile
          340          345          350
Asn Leu Thr Val Ser Leu Ala Ser Thr Asn Leu Asp Tyr Gln Glu Thr
          355          360          365
Phe Glu Leu His Arg Thr Asn Gln Lys Val Leu Gln Thr Ala Ala Ile
          370          375          380
Pro Ser Leu Pro Thr Gly Leu Phe Val Ser Ala Lys Gly Asp Gly Cys
          385          390          395
Cys Leu Met Gln Ile Asp Val Thr Tyr Asn Val Pro Asp Pro Val Ala
          400          405          415
Lys Pro Ala Phe Gln Leu Leu Val Ser Leu Gln Glu Pro Glu Ala Gln
          420          425          430
Gly Arg Pro Pro Pro Met Pro Ala Ser Ala Ala Glu Gly Ser Arg Gly
          435          440          445
Asp Trp Pro Pro Ala Asp Asp Asp Asp Pro Ala Ala Asp Gln His His
          450          455          460
Gln Glu Tyr Lys Val Met Leu Glu Val Cys Thr Arg Trp Leu His Ala
          465          470          475
Gly Ser Ser Asn Met Ala Val Leu Glu Val Pro Leu Leu Ser Gly Phe
          485          490          495
Arg Ala Asp Ile Glu Ser Leu Glu Gln Leu Leu Asp Lys His Met
          500          505          510
Gly Met Lys Arg Tyr Glu Val Ala Gly Arg Arg Val Leu Phe Tyr Phe
          515          520          525
Asp Glu Ile Pro Ser Arg Cys Leu Thr Cys Val Arg Phe Arg Ala Leu
          530          535          540
Arg Glu Cys Val Val Gly Arg Thr Ser Ala Leu Pro Val Ser Val Tyr
          545          550          555
Asp Tyr Tyr Glu Pro Ala Phe Glu Ala Thr Arg Phe Tyr Asn Val Ser
          565          570          575
Thr His Ser Pro Leu Ala Arg Glu Leu Cys Ala Gly Pro Ala Cys Asn
          580          585          590
Glu Val Glu Arg Ala Pro Ala Arg Gly Pro Gly Trp Phe Pro Gly Glu
          595          600          605
Ser Gly Pro Ala Val Ala Pro Glu Glu Gly Ala Ala Ile Ala Arg Cys
          610          615          620
Gly Cys Asp His Asp Cys Gly Ala Gln Gly Asn Pro Val Cys Gly Ser
          625          630          635
Asp Gly Val Val Tyr Ala Ser Ala Cys Arg Leu Arg Glu Ala Ala Cys
          645          650          655
Arg Gln Ala Ala Pro Leu Glu Pro Ala Pro Pro Ser Cys Cys Ala Leu
          660          665          670
Glu Gln Arg Leu Pro Ala Ser Ser Ser Ser Thr Tyr Gly Asp Asp Leu
          675          680          685
Ala Ser Val Ala Pro Gly Pro Leu Gln Gln Asp Val Lys Leu Asn Gly
          690          695          700
Ala Gly Leu Glu Val Glu Asp Ser Asp Pro Glu Pro Glu Gly Glu Ala
          705          710          715
Glu Asp Arg Val Thr Ala Gly Pro Arg Pro Pro Val Ser Ser Gly Asn
          725          730          735
Leu Glu Ser Ser Thr Gln Ser Ala Ser Pro Phe His Arg Trp Gly Gln

```

```

              740              745              750
Thr Pro Ala Pro Gln Arg His Ser Gly Arg Val Val Gly Ala His Arg
              755              760              765
Pro Gly Leu Leu Ser Pro Val Phe Val Tyr Ser Pro Ala Phe Gln Ser
              770              775              780
Gly Gly Glu Glu Gly Leu Trp Met Ser Asn Thr Cys Thr Leu Arg
              785              790              795

```

```

<210> 79
<211> 346
<212> DNA
<213> Homo sapiens

```

```

<400> 79
acgcgttgcc gtcggggaag tggtagatgt aaggtcttcc ctttctcgcg gcggcttctt
60
cgctgatagg cgagaggttc atgccgccaa agcccgggaa acccacgtcg taggcatccg
120
ccgatgcaac ggcataata ctcaaggcca tgaagtcggt ggtgcggctc tggaagtact
180
tcaccgcgcc ttccgacaag cccacgtcct tgagcaggaa gtgcggttag ctggtggccg
240
ccagatactc ggctttttct tcggcggtact tgccccggcag gtaatccttg ggcgcgacgt
300
gcattggcgat caatgcccggt ctgtccgctt ccggnccnnn cnnccnn
346

```

```

<210> 80
<211> 101
<212> PRT
<213> Homo sapiens

```

```

<400> 80
Met His Val Ala Pro Lys Asp Tyr Leu Pro Gly Lys Ser Ala Glu Glu
1      5      10      15
Lys Ala Glu Tyr Leu Ala Ala Thr Ser Tyr Arg Asp Phe Leu Leu Lys
20     25     30
Asp Val Gly Leu Ser Glu Gly Ala Val Lys Tyr Phe Gln Ser Arg Thr
35     40     45
Asn Asp Phe Met Ala Leu Ser Ile Asp Ala Val Ala Ser Ala Asp Ala
50     55     60
Tyr Ser Val Gly Phe Pro Gly Phe Gly Gly Met Asn Leu Ala Pro Ile
65     70     75     80
Ser Glu Glu Ala Ala Glu Met Glu Glu Pro Tyr Ile Tyr His Phe
85     90     95
Pro Asp Gly Asn Ala
100

```

```

<210> 81
<211> 429
<212> DNA
<213> Homo sapiens

```

```

<400> 81

```

```

nngcttttct cncctcgc cctgacatgc tgcagacgtg cgggaggggc attaaaagat
60
tgacttaaat ttgataccag aaatgtcagt gtggcttttc ttgaacagcc gcgtgcgggg
120
cctctatgag tggaaatccag tctcatggcc ccccccattg ctcctgttac cctggaggag
180
gctactccat gaggtcggct ccggtggcac tgcctcgggg ctgcatgtac atgtgtgtgc
240
atgnntgtgt gcatgtgcgt gtgcacgtgt nngtgtgtgc ccgtgtgcat gtgcccggtg
300
gcgtgtctgt gctgtgtgtg tgcgtgcatg tgtgcgtgtc tgtgcgtgnc tgtgtgtgtg
360
cgtgcatgtg tgtgcacgtg tgtgcatgtg cactgtgtgc caccgtgtg catgtccgtg
420
tgcacgcgt
429

```

<210> 82

<211> 79

<212> PRT

<213> Homo sapiens

<400> 82

```

Gly Trp Leu Arg Trp His Cys Leu Gly Ala Ala Cys Thr Cys Val Cys
1          5          10          15
Met Xaa Val Cys Met Cys Val Cys Thr Cys Xaa Cys Val Pro Val Cys
20          25          30
Met Cys Pro Cys Ala Cys Leu Cys Cys Val Cys Ala Cys Met Cys Ala
35          40          45
Cys Leu Cys Val Xaa Val Cys Val Arg Ala Cys Val Cys Thr Cys Val
50          55          60
His Val His Val Cys Ala Pro Val Cys Met Ser Val Cys Thr Arg
65          70          75

```

<210> 83

<211> 411

<212> DNA

<213> Homo sapiens

<400> 83

```

gtgcactttc tcggacgagg gtccttcggc gcccgcgata acgagagggc gcgagattt
60
cgtgcgtctc ttaaaaagcg caaagccttt tataagttcg atgtgctttt ttccaggccc
120
ggaaaagcgc gatccgtaga taaaataggg ccgctttatc gcaaatggct ttatggttac
180
cgtttcgcga tcgatttgca gctgcggaaa aaaaggcctg tggtcgatgc cgttgtggat
240
gatttcgatt ttcgctttgt cgatgccaaag ccctatcaaa tcgttttcgca aaaactcgct
300
cggggcgacg atcttttgca cctttgtaag ccccgcttt atctgcatgc gcatcatgag
360
ctcgctcttt tctttgagca tgacggaaac gatggaattg acgacggcga c
411

```

<210> 84
 <211> 127
 <212> PRT
 <213> Homo sapiens

<400> 84
 Met Leu Lys Glu Lys Ser Asp Arg Met Met Arg Met Gln Ile Lys Arg
 1 5 10 15
 Gly Leu Thr Lys Val Gln Lys Ile Val Ala Ala Ser Glu Phe Leu Arg
 20 25 30
 Asn Asp Leu Ile Gly Leu Gly Ile Asp Lys Ala Lys Ile Glu Ile Ile
 35 40 45
 His Asn Gly Ile Asp His Arg Pro Phe Phe Pro Gln Leu Gln Ile Asp
 50 55 60
 Ala Glu Thr Val Thr Ile Lys Pro Phe Ala Ile Lys Arg Pro Tyr Phe
 65 70 75 80
 Ile Tyr Gly Ser Arg Leu Ser Gly Pro Glu Lys Lys His Ile Glu Leu
 85 90 95
 Ile Lys Ala Phe Ala Leu Phe Lys Glu Arg Thr Lys Ser Pro His Pro
 100 105 110
 Leu Val Ile Ala Gly Ala Glu Gly Pro Ser Ser Glu Glu Val His
 115 120 125

<210> 85
 <211> 333
 <212> DNA
 <213> Homo sapiens

<400> 85
 nnacgcgtgc cgcgcaaggg aaccatgttc gtgtcggtgc gcgaaaccga caagcgcgcg
 60
 atctctgacg cgggtgaaact gctgagttcg ctccggttca aggtgatcgc cacctcgggc
 120
 acccagcgtt tcttgggtgga gaacggagta ccggcggaag agatcaacaa ggtgctggaa
 180
 ggccgcccgc acattgttga cgcaattacc aacggcgagg tgcaactcgt tttcaatacc
 240
 accgaggggc cacaggcgct ggctgacagc cgctcgttgc gacgcgctgc cctcttgcac
 300
 aaagtgccat attacaccac tctttcaggt gca
 333

<210> 86
 <211> 111
 <212> PRT
 <213> Homo sapiens

<400> 86
 Xaa Arg Val Pro Arg Lys Gly Thr Met Phe Val Ser Val Arg Glu Thr
 1 5 10 15
 Asp Lys Ala Arg Ile Leu Asp Ala Val Lys Leu Leu Ser Ser Leu Gly
 20 25 30
 Phe Lys Val Ile Ala Thr Ser Gly Thr Gln Arg Phe Leu Val Glu Asn

ATTORNEY DOCKET NO.: 15966-543

```

      35              40              45
Gly Val Pro Ala Glu Lys Ile Asn Lys Val Leu Glu Gly Arg Pro His
      50              55              60
Ile Val Asp Ala Ile Thr Asn Gly Glu Val Gln Leu Val Phe Asn Thr
      65              70              75
Thr Glu Gly Pro Gln Ala Leu Ala Asp Ser Arg Ser Leu Arg Arg Ala
      85              90              95
Ala Leu Leu His Lys Val Pro Tyr Tyr Thr Leu Ser Gly Ala
      100             105             110

```

<210> 87

<211> 355

<212> DNA

<213> Homo sapiens

<400> 87

```

acgcgtgagg aaatgggggc cgcaggcctg gatcgcaagg tatggcagtg cccggtcgtc
60
ctcctgagcg atgttcattc ggtaggggta cagggtagcg ggcgtactta tggttctccc
120
attgtgcttc gcccggtgac gagtgaggac gccatgactg cggactgggc acgtatccca
180
tatgacgtac tggaaaagat ctcgactcgc attacgaatg cgtgtccgca aatcaaccgg
240
gtgggtactcg atatcacatc taaaccgccg gccaccatcg agtgggaatg agccccgtct
300
caccgtagaac atgacatggc ccgcaccttt cttggggcgg gccatgccgt gttag
355

```

<210> 88

<211> 96

<212> PRT

<213> Homo sapiens

<400> 88

```

Thr Arg Glu Glu Met Gly Ala Ala Gly Leu Asp Arg Lys Val Trp Gln
  1              5              10              15
Cys Pro Val Val Leu Leu Ser Asp Val His Ser Val Gly Val Gln Gly
      20              25              30
Asp Gly Arg Thr Tyr Gly Ser Pro Ile Val Leu Arg Pro Val Thr Ser
      35              40              45
Glu Asp Ala Met Thr Ala Asp Trp Ala Arg Ile Pro Tyr Asp Val Leu
      50              55              60
Glu Lys Ile Ser Thr Arg Ile Thr Asn Ala Cys Pro Gln Ile Asn Arg
      65              70              75              80
Val Val Leu Asp Ile Thr Ser Lys Pro Pro Ala Thr Ile Glu Trp Glu
      85              90              95

```

<210> 89

<211> 351

<212> DNA

<213> Homo sapiens

<400> 89

ATTORNEY DOCKET NO.: 15966-543

nacgcgtcaa caccaggcta cgggtgggtat gatcatgata agggctggga ccgcaggag
 60
 agggcaagag ttatataata gacgggtgta tgggtttaag gtatacatg tgatgagaca
 120
 ggtgattgga taaggccaga gggttgggcg gggttctgcc cctgctgaag cctggtgggg
 180
 cccaggtctg tgatctggga ccggaacaac acatctgctc tgggctctgt ggatgtggcg
 240
 caagccctgg aacagaacca cagcctcaag tccatgcccg tgccactgaa tgacgtaacc
 300
 caggctcatc gcagccggcc agaactcaca actcgagcgg tccatcagat c
 351

<210> 90

<211> 61

<212> PRT

<213> Homo sapiens

<400> 90

Ser	Leu	Val	Gly	Pro	Arg	Ser	Val	Ile	Trp	Asp	Arg	Asn	Asn	Thr	Ser
1			5					10					15		
Ala	Leu	Gly	Leu	Leu	Asp	Val	Ala	Gln	Ala	Leu	Glu	Gln	Asn	His	Ser
			20				25					30			
Leu	Lys	Ser	Ser	Met	Pro	Leu	Pro	Leu	Asn	Asp	Val	Thr	Gln	Ala	His
		35					40					45			
Ser	Arg	Pro	Glu	Leu	Thr	Thr	Arg	Ala	Val	His	Gln	Ile			
	50					55					60				

<210> 91

<211> 327

<212> DNA

<213> Homo sapiens

<400> 91

nggtgggtcc tctcgcagga caggtaatcc tgagacctac gcagctccct ggagctctgc
 60
 acgtogaagc ccagcagggc ctctcgcagg tccctggggc agccagcaca cacaagtc
 120
 cgggaagggc tgtagactcc ctgccagcgg ctttcccggg gaaggcacc acgccagct
 180
 gccctcttga ggtactgctc gggctctggg ggaggggcgc cgtgtccagc acaccctgtg
 240
 tgtgcagtcc tctccctgcc ccaactgccga acgagccctc cacggtgaag ccattgggga
 300
 acgtgacctt gcccttcccc atgaggt
 327

<210> 92

<211> 107

<212> PRT

<213> Homo sapiens

<400> 92

Met	Gly	Lys	Gly	Lys	Val	Thr	Phe	Pro	Asn	Gly	Phe	Thr	Val	Glu	Gly
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

ATTORNEY DOCKET NO.: 15966-543

```

      1             5             10             15
Ser Phe Gly Ser Gly Ala Gly Arg Gly Leu His Thr Gln Gly Val Leu
      20             25             30
Asp Thr Ala Ala Leu Pro Pro Asp Pro Ser Ser Thr Cys Lys Arg Gln
      35             40             45
Leu Gly Val Gly Ala Phe Pro Gly Lys Ala Ala Gly Arg Glu Ser Thr
      50             55             60
Ala Pro Ser Gly Thr Leu Cys Val Leu Ala Ala Pro Gly Thr Cys Arg
      65             70             75             80
Arg Pro Cys Trp Ala Ser Thr Cys Arg Ala Pro Gly Ser Cys Val Gly
      85             90             95
Leu Arg Ile Thr Cys Pro Ala Arg Gly Pro Thr
      100             105

```

<210> 93

<211> 394

<212> DNA

<213> Homo sapiens

<400> 93

```

nccgcgtacg acaagcagta cctcgagggg cggtacgggtg cggacccata cctgagcaac
60
atgctcgaat gggacggcgg acatgagcaa taggccgccca aagcacacgc agaggaaggg
120
ggcgcgtctc gcgccctcgc ggtagaagga ggggtggtcg cgcaagagac gccccgcgcc
180
ttttgtaccg ataatcccg ggcgccacct gatcgtgacc gaggtacga aaacggaacc
240
gtctatttcc gaggtatca ggttgcgtgt caacaaccgt tatcacggcc agtgggtgac
300
aatggaagtt gtcgttacgc gcaagcatac caggggactt ctgcgtcgtg cagtcaactct
360
ggcggaagaa agtgccacag gattcaactca cgta
394

```

<210> 94

<211> 109

<212> PRT

<213> Homo sapiens

<400> 94

```

Met Leu Ala Gly Asn Asp Asn Phe His Cys His Pro Leu Ala Val Ile
      1             5             10             15
Thr Val Val Asp Thr Gln Pro Asp Ser Leu Glu Ile Glu Arg Phe Arg
      20             25             30
Phe Arg Thr Leu Gly His Asp Gln Val Ala Pro Gly Ile Tyr Arg Tyr
      35             40             45
Lys Arg Arg Gly Ala Ser Leu Ala Arg Pro Pro Leu Leu His Pro Arg
      50             55             60
Gly Arg Arg Ala Arg Arg Leu Pro Leu Ala Val Leu Trp Arg Pro Ile
      65             70             75             80
Ala His Val Arg Arg Pro Ile Arg Ala Cys Cys Ser Gly Met Gly Pro
      85             90             95
His Arg Asn Ala Pro Arg Gly Thr Ala Cys Arg Thr Arg

```

100

105

<210> 95
<211> 531
<212> DNA
<213> Homo sapiens

<400> 95
gggtacctctt ccaagtacct tctaaatgaa acactcaaga gagtgctact caggaactt
60
tgcttgatgc ctaaaatgga ctggtcttgg gtgtgtaacc cgggtgaagt tatagcctcc
120
ccaaattgag gtgacagaag gaagacaaga ggtgtaagct ggagaggga ggaagaaat
180
cagtggtctt ggccagcctc tgtgccaccc agtacgacag aggagtgga actggccctc
240
tggtggctctg ctggccata ggcactgcac attgtgccac ctgctcatca cctcctctag
300
tctcacactg agcatcggag tacctgttgt gcagacagga aaactgagga gctctgagag
360
gctgagcatg gagctcacc catgccatag ggtgtggga gagggcacag gaggcctcat
420
ccatggggga aagggttgag gatggacatg ggtggggaga gggcatagac atcccttctc
480
aatctctgtt cccaccacat ttcattagg atgagttagg agatgacagc t
531

<210> 96
<211> 124
<212> PRT
<213> Homo sapiens

<400> 96
Met Arg Pro Pro Val Pro Ser Ser His Thr Leu Trp His Gly Val Ser
1 5 10 15
Ser Met Leu Ser Leu Ser Glu Leu Leu Ser Phe Pro Val Cys Thr Thr
20 25 30
Gly Thr Pro Met Leu Ser Val Arg Leu Glu Glu Val Met Ser Arg Trp
35 40 45
His Asn Val Gln Cys Leu Trp Pro Ser Arg Ala Pro Glu Gly Gln Phe
50 55 60
Pro Leu Leu Cys Arg Thr Gly Trp His Arg Gly Trp Pro Lys Pro Leu
65 70 75 80
Ile Ser Ser Leu Pro Ser Pro Ala Tyr Thr Ser Cys Leu Pro Ser Val
85 90 95
Thr Ser Ile Trp Gly Gly Tyr Asn Phe Thr Gly Val Thr His Pro Arg
100 105 110
Pro Val His Phe Arg Ile Gln Ala Lys Phe Pro Glu
115 120

<210> 97
<211> 405
<212> DNA
<213> Homo sapiens

<400> 97
 ggtcggggcca gtcgaacttc attcccgctt cgagggtctt gctgcggatg ggccgtacgc
 60
 tcgcgggtgcc ttgcgcgcgg gctggtaggt ggagaagccg cgcgagtagc cgcgtagag
 120
 cgacatcgtg tctgagacgt cgaagctcag gccagcttt ggcgtccagg cgcgctcggt
 180
 cggtcgccccc tcttcgcgca attgattcag cgcaatccc gccatcacat gccagcgctt
 240
 gtccagggtc atgaaatcct gggcataggc gcgcgaggag cgcagcgccg aattggacag
 300
 gcgctcgata ttgggcgtga tgtccgaaga cgggaacggg acccgggggg agaagacgtt
 360
 gcccgggaaa agatcccccg acgccatcgt ggtgtcgacc gagat
 405

<210> 98
 <211> 122
 <212> PRT
 <213> Homo sapiens

<400> 98
 Met Ala Ser Gly Asp Leu Phe Pro Gly Asn Val Phe Ser Pro Arg Val
 1 5 10 15
 Pro Phe Pro Ser Ser Asp Ile Thr Pro Asn Ile Glu Arg Leu Ser Asn
 20 25 30
 Ser Pro Leu Arg Ser Ser Arg Ala Tyr Ala Gln Asp Phe Met Thr Leu
 35 40 45
 Asp Lys Arg Trp His Val Met Ala Gly Ile Ala Leu Asn Gln Leu Pro
 50 55 60
 Gln Glu Gly Gly Pro Thr Glu Arg Ala Trp Thr Pro Lys Leu Gly Leu
 65 70 75 80
 Ser Phe Asp Val Ser Asp Thr Met Ser Leu Tyr Gly Ala Tyr Ser Arg
 85 90 95
 Gly Phe Ser Thr Tyr Gln Pro Ala Arg Lys Ala Pro Arg Ala Tyr Gly
 100 105 110
 Pro Ser Ala Ala Arg Pro Ser Lys Arg Glu
 115 120

<210> 99
 <211> 545
 <212> DNA
 <213> Homo sapiens

<400> 99
 acgcgtccgc tcccgatgtc gttgacgagc tgcgtccagc gatgacggta etcaccatc
 60
 tgccccgacg acccagcaaa cgtccccggc tgttctctcat tgaccacgcc gaccggtatg
 120
 tcgatcccat cactcgggat ttgctggaat ccttggttcg cgaagccggc gaggcgcgg
 180
 tgatcttggg tgcccagcgt cgcggctcga tcgattggct etccccacag atcatccaca
 240

acctggccga acaccatltt gagtcgtcct ctggaggtac tcgatgatga ctgaacgttc
 300
 ccattccacg atcagggttaa ggtggccggc ggtggtggtt ctgctcctcg ttccgtgct
 360
 ggtggtcgcc ggattggtcc gggacgacct ggcataccac cgaccgggtg ggcgggtga
 420
 aagcggccgt cgtcaacgag gacaaggccg tcaaggtgcg tggacaactg gttccgatgg
 480
 gccgccaaact caccgccgcc ttgatggact ctggctcgca caccactgat ggccacaccg
 540
 tcgac
 545

<210> 100

<211> 101

<212> PRT

<213> Homo sapiens

<400> 100

Met Gly Thr Phe Ser His His Arg Val Pro Pro Glu Asp Asp Ser Lys
 1 5 10 15
 Trp Cys Ser Ala Arg Leu Trp Met Ile Cys Gly Glu Ser Gln Ser Met
 20 25 30
 Arg Pro Arg Arg Trp Ala Pro Lys Ile Thr Ala Ala Ser Pro Ala Ser
 35 40 45
 Arg Thr Arg Asp Ser Ser Lys Ser Arg Val Met Gly Ser Thr Ile Arg
 50 55 60
 Ser Ala Trp Ser Met Arg Asn Ser Arg Gly Arg Leu Leu Gly Arg Arg
 65 70 75 80
 Gly Arg Trp Val Ser Thr Val Ile Ala Glu Arg Ser Ser Ser Thr Thr
 85 90 95
 Ser Gly Ala Asp Ala
 100

<210> 101

<211> 619

<212> DNA

<213> Homo sapiens

<400> 101

ngcgcgccac agaagtga aaagtacggtc tcaagccgtt gggagcggtt ttctgtctcg
 60
 tcattgcacga tccggacttc gatccgatcc ccattggtgaa caaggagctt gacgccttcg
 120
 aagctgcggc ggggtactat ctcatcctcg ccacggatcc cggacgcaag ggatacacga
 180
 ccgcccgtcc tcacgaggcc ggcggaaaac gctattacca acctggacca gatccgcgaa
 240
 gtctgcgcca gccgcaactg caccgcctgt ctacaccccc attggggaac gatgggtccag
 300
 aaccgtgaac aagtgatccg cgtgctcgag aactcctcga tcgggctgtg cctggacact
 360
 ggtcatctgg cctgtggtgg taccgatgct gttgagctgg tgcgtaagta cgccaaccgc
 420

agtggttggg gcgaagaccc c
321

<210> 104
<211> 107
<212> PRT
<213> Homo sapiens

<400> 104
Xaa His Gly Gly Arg Gln Gln Pro Cys Gly Gly Asp Pro Ser His Trp
1 5 10 15
Gly Glu Thr Pro Ala Ile Gly Gly Arg Pro Leu Pro Trp Gly Lys Thr
20 25 30
Pro Ala Met Gly Gln Thr Pro Ala Thr Gly Gly Arg Pro Leu Pro Leu
35 40 45
Gly Gly Asp Pro Ser His Trp Gly Glu Thr Pro Ala Met Gly Lys Asp
50 55 60
Pro Cys His Trp Gly Arg Xaa Pro Ala Ile Gly Gly Asp Pro Cys Arg
65 70 75 80
Trp Gly Glu Ile Pro Ala Val Gly Gly Xaa Pro Pro Val Gly Glu
85 90 95
Asp Pro Cys Arg Ser Gly Trp Gly Glu Asp Pro
100 105

<210> 105
<211> 344
<212> DNA
<213> Homo sapiens

<400> 105
nnmtctctcc gaccgcgtcc agatccaccg tggcccgcat gaaccagtcg ttgttcctt
60
ccgggtcaac gaggggttgg cgacgggtcc actcgtggc gcccggggtg atgtgcaaca
120
gggcgggccc gcgcgcggcc gggcctgatt ccagcctctc gtgctcgtcc cagtaccact
180
ccagcgcac ccccccagcg tcggcctccc agccgtggtc gccgtcagc gccccaggg
240
cttcaatgtc gtcacggcg gccagttcca cccggcgaa catctcgtt cggaccatga
300
cccgaaggc gcgggaattc tcggtcagtt tcggcggtgc cggc
344

<210> 106
<211> 62
<212> PRT
<213> Homo sapiens

<400> 106
Cys Ala Thr Gly Arg Ala Arg Ala Pro Gly Leu Ile Pro Ala Ser
1 5 10 15
Arg Ala Arg Pro Ser Thr His Pro Ala His Arg Pro Ser Gly Arg His
20 25 30
Pro Ser Arg Gly Arg Arg Arg Ala Pro Gly Pro Gln Cys Arg His

	35		40		45
Arg	Arg	Pro	Val	Pro	Pro
	50		55		60

<210> 107

<211> 549

<212> DNA

<213> Homo sapiens

<400> 107

acgcgttcc t cggacgtgc tagtttgagc gcacagggcat tgggcggggc tcacggggga
60
gccgcttaat aaccgaccaa catgaaactc aaggggctgcc ccttcctagc ggggaccctg
120
cacagacccg aaaataagggt gttttgctct gccctcctca gttcacgtgg gcaccttgga
180
acactgaaga aggcattttc cgaactcact gtccctacgga cttattctcc gcaactgtttt
240
cgccctcctc gccctgttct cgtgactgac aggagcagggt gtcacaagca ggcagccgga
300
gagctctgct cacctggaaa agcatttttg ttagctttaa atgtgaaggc ctcaggcagt
360
ggcctgttgt cctcctccac atgcgccccat cttcactctt tcatgtgact ggcctgtttt
420
tgaagggaag gccctgtgca cccttggtta ggcaggttat gttctgcacc gaaaatgggc
480
tgccctctg cattggatgg ctagctctta ggttggttta ttttagcaaa taagcggttac
540
agggtaggc
549

<210> 108

<211> 108

<212> PRT

<213> Homo sapiens

<400> 108

Met	Lys	Leu	Lys	Gly	Cys	Pro	Phe	Leu	Ala	Gly	Thr	Leu	His	Arg	Pro
1				5					10					15	
Glu	Asn	Lys	Gly	Phe	Cys	Ser	Ala	Leu	Leu	Ser	Ser	Arg	Gly	His	Leu
			20					25					30		
Gly	Thr	Leu	Lys	Lys	Ala	Phe	Ser	Glu	Leu	Thr	Val	Leu	Arg	Thr	Tyr
			35				40					45			
Ser	Pro	His	Cys	Phe	Arg	Leu	Leu	Arg	Pro	Val	Leu	Val	Thr	Asp	Arg
			50			55				60					
Ser	Arg	Gly	His	Lys	Gln	Ala	Ala	Arg	Glu	Leu	Cys	Ser	Pro	Gly	Lys
65				70					75					80	
Ala	Phe	Leu	Cys	Ser	Leu	Asn	Val	Lys	Ala	Ser	Gly	Ser	Gly	Leu	Leu
			85					90						95	
Ser	Ser	Ser	Thr	Cys	Ala	His	Leu	His	Ser	Phe	Met				
			100				105								

<210> 109

<211> 748

<212> DNA

<213> Homo sapiens

<400> 109

```

nngaattcag atttactttt tgcatttcct tgaatataaa ataggcagta aactaaccta
60
aagtcagaag acctatttgt tctggtacct attctgctgt taataaatg ataaatgagt
120
gatattgggg taagttacca atttacttta cagcccttaa gtaataatc tgctttcctc
180
agcatcatag accttttgag aggattaatt aagcgcttaa aaaacctgta gactctatta
240
cagtcagtga aaggaataat tctctttaca aagtaaatgc agttgtttta ttttagacaa
300
gagtggttcta aacttcgtga agagttaagg cttcaacatg aagaggataa gaagtcagca
360
atgtctcaac ttttgagtt gaaagatcga gaaaaaatg cagcaagaga ttcattggcag
420
aagaaagtag aagatctctt aaaccagatt tccttgctga aacagaatct ggagatacag
480
ctttccagct ctcagacttc ttgcaacaa ctgcaagccc agtttacgca agaacgacag
540
cggcttacgc aagagcttga agaattagag gagcaacatc agcaaaagaca caaatcatta
600
aaagaagcac atgtccttgc atttcaaact atggaagagg aaaagggaaa ggagcaaaaa
660
gctcttgaaa atcatttaca acagaagcat tctgcagagc ttcaatcact aaaagatgca
720
cacagagagt caatggaggg cttccgga
748

```

<210> 110

<211> 157

<212> PRT

<213> Homo sapiens

<400> 110

```

Met Gln Leu Phe Tyr Phe Arg Gln Glu Cys Ser Lys Leu Arg Glu Glu
1           5           10           15
Leu Arg Leu Gln His Glu Glu Asp Lys Lys Ser Ala Met Ser Gln Leu
20           25           30
Leu Gln Leu Lys Asp Arg Glu Lys Asn Ala Ala Arg Asp Ser Trp Gln
35           40           45
Lys Lys Val Glu Asp Leu Leu Asn Gln Ile Ser Leu Leu Lys Gln Asn
50           55           60
Leu Glu Ile Gln Leu Ser Gln Ser Gln Thr Ser Leu Gln Gln Leu Gln
65           70           75           80
Ala Gln Phe Thr Gln Glu Arg Gln Arg Leu Thr Gln Glu Leu Glu Glu
85           90           95
Leu Glu Glu Gln His Gln Gln Arg His Lys Ser Leu Lys Glu Ala His
100          105          110
Val Leu Ala Phe Gln Thr Met Glu Glu Glu Lys Glu Lys Glu Gln Arg
115          120          125
Ala Leu Glu Asn His Leu Gln Gln Lys His Ser Ala Glu Leu Gln Ser

```


130 135 140
 Leu Lys Asp Ala His Arg Glu Ser Met Glu Gly Phe Arg
 145 150 155

<210> 111
 <211> 429
 <212> DNA
 <213> Homo sapiens

<400> 111
 gcgcgccag agagtcaca atgggtgtcag gatatgggag ccaccggaat cattaaccat
 60
 cgtgaacgcg cgcgccgca ggtcaacgac ttccgggatca ccgggttcga eggcattctc
 120
 tcggcttatn nacgccacca gcatncgact ttggctgaga tcatcgacc gttcggacat
 180
 ctggctatga tcgacggaac cgactcattc gatctcatgg ccttcaagtc aaagtcgtta
 240
 acggtgacca gcgagtcgat gttcagccgt ccacagttcg cgacgcccga cgtcgccgaa
 300
 caaggccggg cactggccag catcgccgac ctctcgaga aggggcagat ccgtccgacg
 360
 atgacccgcc atatcgaggg tctgacaacc cagcatgtgc gtgaggccac cgcagccgtc
 420
 gagtccggc
 429

<210> 112
 <211> 143
 <212> PRT
 <213> Homo sapiens

<400> 112
 Ala Arg Pro Glu Ser Ala Gln Trp Cys Gln Asp Met Gly Ala Thr Gly
 1 5 10 15
 Ile Ile Asn His Arg Glu Pro Leu Ala Pro Gln Val Asn Asp Phe Gly
 20 25 30
 Ile Thr Gly Phe Asp Gly Ile Leu Ser Ala Tyr Xaa Arg His Gln His
 35 40 45
 Xaa Thr Leu Ala Glu Ile Ile Ala Pro Phe Gly His Leu Val Met Ile
 50 55 60
 Asp Gly Thr Asp Ser Phe Asp Leu Met Ala Phe Lys Ser Lys Ser Leu
 65 70 75 80
 Thr Val Thr Ser Glu Ser Met Phe Ser Arg Pro Gln Phe Ala Thr Pro
 85 90 95
 Asp Val Ala Glu Gln Gly Arg Ala Leu Ala Ser Ile Ala Asp Leu Val
 100 105 110
 Glu Lys Gly Gln Ile Arg Pro Thr Met Thr Arg His Ile Glu Gly Leu
 115 120 125
 Thr Thr Gln His Val Arg Glu Ala Thr Ala Ala Val Glu Ser Gly
 130 135 140

<210> 113
 <211> 382

<212> DNA

<213> Homo sapiens

<400> 113

gtgcaactgga cggatgctgg ggaactgagg cctgcccttc tccactcctg ctgcctctgt
 60
 gtggaccgccc tgccacctgg ctccaacgat gtggacgctc tgtgccgggc gctgtcagct
 120
 gtccacagccc ccacctctctg ccagctggcg tgccggccagg atgggcagct gaagggcttc
 180
 gcggtgctgg agtatgagac ggctgagatg gcggaggagg cacagcagca ggcggacggc
 240
 ctgtccctgg ggggcagcca cctgcgagtc tcttctctgc ccctggggac ccccgggcgc
 300
 agtatgctgg ccgctctcat cgttgcccag gccacggccc tcaatcgggg gcaggagtc
 360
 ctccccgagc ccaacatcct gc
 382

<210> 114

<211> 125

<212> PRT

<213> Homo sapiens

<400> 114

Met	Leu	Gly	Ser	Gly	Arg	Thr	Pro	Cys	Pro	Arg	Leu	Arg	Ala	Val	Ala
1				5					10				15		
Trp	Ala	Thr	Met	Arg	Ala	Ala	Ser	Ile	Leu	Arg	Pro	Gly	Val	Pro	Gly
			20					25				30			
Ala	Gln	Lys	Glu	Thr	Arg	Arg	Trp	Leu	Pro	Pro	Arg	Asp	Arg	Pro	Ser
		35					40					45			
Ala	Cys	Cys	Cys	Ala	Ser	Ser	Ala	Ile	Ser	Ala	Val	Ser	Tyr	Ser	Ser
		50				55					60				
Thr	Ala	Lys	Pro	Phe	Ser	Cys	Pro	Ser	Trp	Pro	His	Ala	Ser	Trp	Gln
65				70					75					80	
Lys	Val	Gly	Leu	Trp	Thr	Ala	Asp	Ser	Ala	Arg	His	Arg	Ala	Ser	Thr
			85					90					95		
Ser	Leu	Lys	Pro	Gly	Gly	Arg	Arg	Ser	Thr	Gln	Arg	Gln	Gln	Glu	Trp
			100					105					110		
Arg	Arg	Ala	Gly	Leu	Ser	Ser	Pro	Ala	Ser	Val	Gln	Cys			
		115					120					125			

<210> 115

<211> 4798

<212> DNA

<213> Homo sapiens

<400> 115

ctccctccca gtctgtctta agtactgatg ccagagaggt ctgcctaaagc tcagtgtctgt
 60
 catcttcactc cctgttttga gccatgagtg gccccctgtt gccttcaaga ctgaagccat
 120
 ccccatcctt cctccaccaca acctagaggc ttgtcttcgt aaatgctggc cctttccttc
 180

atgggctcca ccctctgagt gtgtcatgaa ctctccact tccttgccctt gggtcgtatt
240
gggtccctcg cccgaggcca gagatttgga cgagcccttc tcctccatct tcacagcttc
300
catttaccat tagatgccag gccagacag ttgatgaaaa atgtaagaaa cggaccaag
360
tataagaaaa agctagatct gtccatttta tttctagttt ttggaagaga tgtggatata
420
gagaaatatt actctactat tgaaaaaaat atatatcagt ctgattataa atgtcaactc
480
atactaagt aatatcaggg aagcagtggtg gagaaggatg tggcaaacgt gagattgctg
540
tttccataga aagagggcag ctaatgctta tctccagacc attgttatgc agtaatgtag
600
actccatttc accgatttta attcttaagg agaaactgga aacttgtatt attatttgat
660
tattactttt ggcagttaat tttaaaaatc accgttttag ccaaaacaac atgacatttg
720
ctggggtac aagggaactct tagcgactct tagcttgac ttacctttct aatctcaggg
780
ctttttcaa agcccttctc tgaattccct gaagatatgc tagcttgaag gctcactgca
840
tactctcaac tcccagcca ctccctctgt tcctcttcat accagttaat actttgatgt
900
tattttttca gtctatttta atcttttaga ctggagggtc ctgagggtg gggcctaatt
960
cttgttttta tcttcccat agcaattatc aagtattctg ctcccttttg cacctaagaa
1020
atattagttg gggtgataac tgaatggctg tcctcttgcc ttcgttaactg ctttttgga
1080
tgttattttc ttgacctct attccaccg tgtgcagtgg agaggatggg tgttgtaatt
1140
tacagaggaa ggaattaaact gagacgcaac aggatgatgg gaacttctca aggtcatgtg
1200
gcaaggaaat caaggaattg gggtttgaac ccaagtcgtc tcagttccat tccattgtcc
1260
cttaaccttc gccacctctc accttcttcc ctttacctt tttctgtagc tgaacggaaa
1320
ccaccgctct ttaacatgaa tgcgatgagt gccttatacc acattgcaca gaacgaatcc
1380
cccgctgtcc agtcaggaca ctgggtctgag tacttccgga attttgtcga ctctgtctt
1440
cagaaaaatcc ctcaagacag accaacctca gaggttctcc tgaagcacgg ctttgtgtc
1500
cgggagcggc caccacagc catcatggac ctgatccaga ggaccaagga tgcgtgcgg
1560
gagctggaca acctgcagta ccgcaagatg aagaagatcc tgttccaaga ggcaccaaac
1620
ggcctgggtg ccgaggcccc agaggaggaa gaggaggccg agccctacat gcaccgggcc
1680
gggactctga ccagcctcga gagtagccac tcagtgccca gcatgtccat cagcgcctcc
1740
agccagagca gctccgtcaa cagcctagca gatgcctcag acaacaggga agaggaggag
1800

gaggaggagg aagaggagga ggaggaaagaa ggccctgaag cccgggagat ggccatgatg
1860
caggaggggg agcacacagt cacctctcac agctccatta tccaccggct gccgggctct
1920
gacaacctat atgatgacct ctaccagcca gagataacct ccagccctct ccaggccgct
1980
gcagcccccag ctccccacttc caccacctct tccgcccgcc gccgggacct ctgccgtaac
2040
cgagaccact ttgccaccat ccgaaccgcc tccttggtca gccgtcagat ccaggagcat
2100
gagcaggact ctgcgctgcg ggagcagctg agcgggtata agcggatgag acgacagcac
2160
cagaagcagc tgctggccct ggagtcacgg ctgaggggtg aacgggagga gcacagtgc
2220
cggctgcagc gggagcttga ggcgacggcg gctggctttg gggcagaggc agaaaagctg
2280
gcccgcgccg accaggccat aggtgagaag gaggcacgag ctgccaggcg cggaggcggg
2340
aagttccagc agcacatcct tgggcagcag aagaaggagc tggctgccct gctggaggca
2400
cagaagcggg cctacaaact tcgcaaggaa cagctgaagg aggagctcca ggagaacccc
2460
agcaactcca agcgggagaa ggccgagtggt ctgctgcggc agaaggagca gctccagcag
2520
tgccaggcgg aggaggaagc agggctgctg cgccggcagc gccagtactt tgagctgcag
2580
tgtcgccagt acaagcgcaa gatgttgctg gctcggcaca gcctggacca ggacctgctg
2640
cgggaggacc tgaacaagaa gcagaccagc aaggacttgg agtgtgcact gctgcttcgg
2700
cagcacgagg ccacgcggga gctggagctg cgccagctcc aggcctgca gcgcacgcgg
2760
gctgagctca cccgcctgca gcaccagacg gagctgggca accagctgga gtacaacaag
2820
cggcgtgagc aagagttgag gcagaagcat gcggcccagg ttgccagca gcccaagagc
2880
ctcaaagtac gtgcaggcca gcgccccccg ggccctccac tccccattcc tggggctctg
2940
ggcccaccca acacaggcac ccctatagaa cagcagccct gctcacctgg ccaggaggca
3000
gtcctggacc aaagaatgct tggcgaggag gaggaagcag ttggagagag aaggattctg
3060
ggaaaggaag gggccacttt ggagcccaag cagcagagga ttctggggga agaatacagg
3120
gcccttagtc ccagtcaca aaaacatggg agcctgggtt atgaggaagt ttggggctctg
3180
cctgaggaga tagaggagct taggggtgcc tccttgtag cccaggagag gagcatgtt
3240
ggccaggagg aggctgggac gtggagcttg tgggggaagg aggatgagag tcttctggat
3300
gaggagtttg agcttggctg ggtccagggc ccagcactga ctcccgtccc tgaggaggag
3360
gaagaagagg aagagggggc tccgattggg acccctaggg atcctggaga tggttgtcct
3420

tcccccgaca tccctcctga accccctcca acacacctga ggccctgccc tggcagccag
 3480
 ctccctggac tctctgtcca tggcctcctg gcccgctctc cctttgcagt ggggtctctc
 3540
 tctggcctcc tggccctcct gctgctgctg ctgcttccat tgcctggcagc ccagggtggg
 3600
 ggtggcctgc aggcagcgct gctggccctt gaggtggggc tgggtgggtct gggggcctcc
 3660
 tacctgctcc tttgtacagc cctgcacctg cctccagtc ttttctact cctggccacg
 3720
 ggtaccgcac tgggggcccgt cctgggcctg agctggcgcc gaggcctcat ggtgttccc
 3780
 ctgggccttg gagctgcctg gctcttagct tggccaggcc tagctctacc tctggtggct
 3840
 atggcagcgg ggggcagatg ggtgcggcag caaggccccc ggggtgcgcg gggcatactt
 3900
 cgaactctgtg tgcgggttct gctgcgcctg tccccatgg ccttccgggc cctgcagggc
 3960
 tgtggggctg tgggggaccg ggggctgttt gcactgtacc ccaaaaccaa caaggatggc
 4020
 ttccgcagcc gcctgcccgt cctggggccc cggcggcgta atccccgcac caccacaac
 4080
 ccattagctc tgttgccaag ggtctgggtc ctgtgcaagg gctggaactg gctctggca
 4140
 cgggccagcc aggggtttagc atcccacttg ccccctggg ccattccacac actggccagc
 4200
 tggggcctgc ttcggggtga acggcccacc cgaatccccc ggctactacc acgcagccag
 4260
 gcgcagctgg gggccctgc ctcccgcag ccactgccag ggaactctagc cgggcggagg
 4320
 tcacgcaccc gccagtcctg gcccttgccc cctggaggt agctgactcc agcccttcca
 4380
 gcccatactc agagcattga gcactttatc tcccacgact cagtgaagtt tctccagtcc
 4440
 ctagtctctc cttttcaccc accttctca gtttgctcac ttacccagc ccagccctt
 4500
 cggacctcta gacaggcagc ctctcagct gtggagtcca gcagtcactc tgtgttctcc
 4560
 tggcgctcct cccctaagtt attgctgttc gcccgctgtg tgtgctcacc ctacccctca
 4620
 tgaactcagc cctggggcca ggggtggtgg aggggtggaa gagtcatgtt ttttttctcc
 4680
 tctttgattt tgtttttctg tctccttcc aacctgtccc cttcccccca ccaaaaaaag
 4740
 aaaaagacaa acacaaataa aatatctgag cggaactgtg aaaaaaaaaa aaaaaaag
 4798

<210> 116

<211> 1062

<212> PRT

<213> Homo sapiens

<400> 116

Met Met Gly Thr Ser Gln Gly His Val Ala Arg Lys Ser Arg Asn Trp

1	5	10	15
Gly Leu Asn Pro Ser Arg Leu Ser Ser Ile Pro Leu Ser Ser Thr Pro	20	25	30
Cys His Leu Ser Pro Ser Ser Leu Ser Pro Phe Ser Val Ala Glu Arg	35	40	45
Lys Pro Pro Leu Phe Asn Met Asn Ala Met Ser Ala Leu Tyr His Ile	50	55	60
Ala Gln Asn Glu Ser Pro Val Leu Gln Ser Gly His Trp Ser Glu Tyr	65	70	75
Phe Arg Asn Phe Val Asp Ser Cys Leu Gln Lys Ile Pro Gln Asp Arg	85	90	95
Pro Thr Ser Glu Val Leu Leu Lys His Arg Phe Val Leu Arg Glu Arg	100	105	110
Pro Pro Thr Val Ile Met Asp Leu Ile Gln Arg Thr Lys Asp Ala Val	115	120	125
Arg Glu Leu Asp Asn Leu Gln Tyr Arg Lys Met Lys Lys Ile Leu Phe	130	135	140
Gln Glu Ala Pro Asn Gly Pro Gly Ala Glu Ala Pro Glu Glu Glu Glu	145	150	155
Glu Ala Glu Pro Tyr Met His Arg Ala Gly Thr Leu Thr Ser Leu Glu	165	170	175
Ser Ser His Ser Val Pro Ser Met Ser Ile Ser Ala Ser Ser Gln Ser	180	185	190
Ser Ser Val Asn Ser Leu Ala Asp Ala Ser Asp Asn Glu Glu Glu Glu	195	200	205
Glu Glu Glu Glu Glu Glu Glu Glu Glu Gly Pro Glu Ala Arg	210	215	220
Glu Met Ala Met Met Gln Glu Gly Glu His Thr Val Thr Ser His Ser	225	230	235
Ser Ile Ile His Arg Leu Pro Gly Ser Asp Asn Leu Tyr Asp Asp Pro	245	250	255
Tyr Gln Pro Glu Ile Thr Pro Ser Pro Leu Gln Pro Pro Ala Pro	260	265	270
Ala Pro Thr Ser Thr Thr Ser Ser Ala Arg Arg Arg Ala Tyr Cys Arg	275	280	285
Asn Arg Asp His Phe Ala Thr Ile Arg Thr Ala Ser Leu Val Ser Arg	290	295	300
Gln Ile Gln Glu His Glu Gln Asp Ser Ala Leu Arg Glu Gln Leu Ser	305	310	315
Gly Tyr Lys Arg Met Arg Arg Gln His Gln Lys Gln Leu Leu Ala Leu	325	330	335
Glu Ser Arg Leu Arg Gly Glu Arg Glu Glu His Ser Ala Arg Leu Gln	340	345	350
Arg Glu Leu Glu Ala Gln Arg Ala Gly Phe Gly Ala Glu Ala Glu Lys	355	360	365
Leu Ala Arg Arg His Gln Ala Ile Gly Glu Lys Glu Ala Arg Ala Ala	370	375	380
Gln Ala Glu Glu Arg Lys Phe Gln Gln His Ile Leu Gly Gln Gln Lys	385	390	395
Lys Glu Leu Ala Ala Leu Leu Glu Ala Gln Lys Arg Thr Tyr Lys Leu	405	410	415
Arg Lys Glu Gln Leu Lys Glu Glu Leu Gln Glu Asn Pro Ser Thr Pro	420	425	430
Lys Arg Glu Lys Ala Glu Trp Leu Leu Arg Gln Lys Glu Gln Leu Gln			

435 440 445
 Gln Cys Gln Ala Glu Glu Glu Ala Gly Leu Leu Arg Arg Gln Arg Gln
 450 455 460
 Tyr Phe Glu Leu Gln Cys Arg Gln Tyr Lys Arg Lys Met Leu Leu Ala
 465 470 475 480
 Arg His Ser Leu Asp Gln Asp Leu Leu Arg Glu Asp Leu Asn Lys Lys
 485 490 495
 Gln Thr Gln Lys Asp Leu Glu Cys Ala Leu Leu Arg Gln His Glu
 500 505 510
 Ala Thr Arg Glu Leu Glu Leu Arg Gln Leu Gln Ala Val Gln Arg Thr
 515 520 525
 Arg Ala Glu Leu Thr Arg Leu Gln His Gln Thr Glu Leu Gly Asn Gln
 530 535 540
 Leu Glu Tyr Asn Lys Arg Arg Glu Gln Glu Leu Arg Gln Lys His Ala
 545 550 555 560
 Ala Gln Val Arg Gln Gln Pro Lys Ser Leu Lys Val Arg Ala Gly Gln
 565 570 575
 Arg Pro Pro Gly Leu Pro Leu Pro Ile Pro Gly Ala Leu Gly Pro Pro
 580 585 590
 Asn Thr Gly Thr Pro Ile Glu Gln Gln Pro Cys Ser Pro Gly Gln Glu
 595 600 605
 Ala Val Leu Asp Gln Arg Met Leu Gly Glu Glu Glu Glu Ala Val Gly
 610 615 620
 Glu Arg Arg Ile Leu Gly Lys Glu Gly Ala Thr Leu Glu Pro Lys Gln
 625 630 635 640
 Gln Arg Ile Leu Gly Glu Glu Ser Gly Ala Pro Ser Pro Ser Pro Gln
 645 650 655
 Lys His Gly Ser Leu Val Asp Glu Glu Val Trp Gly Leu Pro Glu Glu
 660 665 670
 Ile Glu Glu Leu Arg Val Pro Ser Leu Val Pro Gln Glu Arg Ser Ile
 675 680 685
 Val Gly Gln Glu Glu Ala Gly Thr Trp Ser Leu Trp Gly Lys Glu Asp
 690 695 700
 Glu Ser Leu Leu Asp Glu Glu Phe Glu Leu Gly Trp Val Gln Gly Pro
 705 710 715 720
 Ala Leu Thr Pro Val Pro Glu Glu Glu Glu Glu Glu Glu Ala
 725 730 735
 Pro Ile Gly Thr Pro Arg Asp Pro Gly Asp Gly Cys Pro Ser Pro Asp
 740 745 750
 Ile Pro Pro Glu Pro Pro Pro Thr His Leu Arg Pro Cys Pro Ala Ser
 755 760 765
 Gln Leu Pro Gly Leu Leu Ser His Gly Leu Leu Ala Gly Leu Ser Phe
 770 775 780
 Ala Val Gly Ser Ser Ser Gly Leu Leu Pro Leu Leu Leu Leu Leu
 785 790 795 800
 Leu Pro Leu Leu Ala Ala Gln Gly Gly Gly Leu Gln Ala Ala Leu
 805 810 815
 Leu Ala Leu Glu Val Gly Leu Val Gly Leu Gly Ala Ser Tyr Leu Leu
 820 825 830
 Leu Cys Thr Ala Leu His Leu Pro Ser Ser Leu Phe Leu Leu Ala
 835 840 845
 Gln Gly Thr Ala Leu Gly Ala Val Leu Gly Leu Ser Trp Arg Arg Gly
 850 855 860
 Leu Met Gly Val Pro Leu Gly Leu Gly Ala Trp Leu Leu Ala Trp

```

865              870              875              880
Pro Gly Leu Ala Leu Pro Leu Val Ala Met Ala Ala Gly Gly Arg Trp
              885              890              895
Val Arg Gln Gln Gly Pro Arg Val Arg Arg Gly Ile Ser Arg Leu Trp
              900              905              910
Leu Arg Val Leu Leu Arg Leu Ser Pro Met Ala Phe Arg Ala Leu Gln
              915              920              925
Gly Cys Gly Ala Val Gly Asp Arg Gly Leu Phe Ala Leu Tyr Pro Lys
              930              935              940
Thr Asn Lys Asp Gly Phe Arg Ser Arg Leu Pro Val Pro Gly Pro Arg
              945              950              955
Arg Arg Asn Pro Arg Thr Thr Gln His Pro Leu Ala Leu Leu Ala Ser
              965              970              975
Val Trp Val Leu Cys Lys Gly Trp Asn Trp Arg Leu Ala Arg Ala Ser
              980              985              990
Gln Gly Leu Ala Ser His Leu Pro Pro Trp Ala Ile His Thr Leu Ala
              995              1000              1005
Ser Trp Gly Leu Leu Arg Gly Glu Arg Pro Thr Arg Ile Pro Arg Leu
              1010              1015              1020
Leu Pro Arg Ser Gln Arg Gln Leu Gly Pro Pro Ala Ser Arg Gln Pro
              1025              1030              1035
Leu Pro Gly Thr Leu Ala Gly Arg Arg Ser Arg Thr Arg Gln Ser Arg
              1045              1050              1055
Ala Leu Pro Pro Trp Arg
              1060

```

<210> 117

<211> 471

<212> DNA

<213> Homo sapiens

<400> 117

```

naccggttga cgatctgtct ggctggtgta gtgatctgcg ctgtgggtgt cgtcgatgac
60
ctgcgcgacc ttctgcctt ggccaaggca gctggccagg tattagcggc cggcacgtc
120
gtcacgggcg gagtgcgaat gttttggatc ccgctgccga actccatcat tgctttgggg
180
acgcctactt cgatcttggg gacggtgttc ttcattgtgt tgtgcgccaa tgcgggtgaat
240
ttcattgatg gacttgacgg cctggcatcc ggtgtggtgg ccacggggtc cttggctttc
300
ttctcataca cctaectgct ggctcacgaa caggactttg ttgttgcgac gactaccagt
360
ctcattacgg ctgcgacggc gggcgctgt ctcggttttt tgccccacaa ctggcatccg
420
gcgaggatgt tcatgggtga ttccggagct ctgctacttg gcttattgct a
471

```

<210> 118

<211> 157

<212> PRT

<213> Homo sapiens


```

<400> 118
Xaa Ala Leu Thr Ile Cys Leu Ala Gly Val Val Ile Cys Ala Val Gly
 1           5           10           15
Val Val Asp Asp Leu Leu Asp Leu Pro Ala Leu Ala Lys Ala Ala Gly
          20           25           30
Gln Val Leu Ala Ala Gly Ile Val Val Thr Gly Gly Val Arg Met Phe
          35           40           45
Trp Ile Pro Leu Pro Asn Ser Ile Ile Ala Leu Gly Thr Pro Thr Ser
          50           55           60
Ile Leu Val Thr Val Phe Phe Ile Val Leu Cys Ala Asn Ala Val Asn
          65           70           75           80
Phe Ile Asp Gly Leu Asp Gly Leu Ala Ser Gly Val Val Ala Ile Gly
          85           90           95
Ser Leu Ala Phe Phe Ser Tyr Thr Tyr Leu Leu Ala His Glu Gln Asp
          100          105          110
Phe Val Val Ala Thr Thr Thr Ser Leu Ile Thr Ala Ala Thr Ala Gly
          115          120          125
Ala Cys Leu Gly Phe Leu Pro His Asn Trp His Pro Ala Arg Met Phe
          130          135          140
Met Gly Asp Ser Gly Ala Leu Leu Leu Gly Leu Leu
          145          150          155

```

```

<210> 119
<211> 302
<212> DNA
<213> Homo sapiens

```

```

<400> 119
ntcaaacatg agcagtcgtg gcggccgagg ccgcgggtggc tattatcgcg agctttatgg
 60
tagccgaggt cgaggcagta aatctaataa aacttttcgca aaaaattcgg atgtctactc
 120
tcagaaaaag actcgaacag tacgaggcac ctccgaagat ttagcacgat cgtccataa
 180
gcttcatatg cgcccgatcc ctgcgtatca tgacattgag ggtatgtggg ctttcccgac
 240
ctttactttt tatctggatc atgcacaagc agaccatac gctgccccaa ataaggcacg
 300
cn
 302

```

```

<210> 120
<211> 98
<212> PRT
<213> Homo sapiens

```

```

<400> 120
Met Ser Ser Arg Gly Gly Arg Gly Arg Gly Tyr Tyr Arg Glu Leu
 1           5           10           15
Tyr Gly Ser Arg Gly Arg Gly Ser Lys Ser Asn Glu Thr Phe Ala Lys
          20           25           30
Asn Ser Asp Val Tyr Ser Gln Lys Lys Thr Arg Thr Val Arg Gly Thr
          35           40           45
Ser Glu Asp Leu Ala Arg Ser Leu His Lys Leu His Met Arg Pro Tyr

```

```

      50              55              60
Pro Ala Tyr His Asp Ile Glu Gly Met Trp Ala Phe Pro Ala Phe Thr
65              70              75              80
Phe Tyr Leu Asp His Ala Gln Ala Asp Pro Tyr Ala Ala Pro Asn Lys
      85              90              95
Ala Arg

```

```

<210> 121
<211> 318
<212> DNA
<213> Homo sapiens

```

```

<400> 121
ngcatggggg gccctgggac cgcacttggt cccctttttt ttttagggaa aaaattgagc
60
cctaaaggat ttgccgcatt acaggaaagt tttttggtta gtttgggggtt gtttctgtgc
120
tgtgtgagaa ggagtagaag cagctccagt agagtgggcc ttttcatttt tatccagagg
180
aaattttgtag gctgtggcta ttacttcctt tttttttttt tttttttttt ttttagagaca
240
gagctctgnt ctgtcgccag gctggagtga agtggcacga tctcagctca ctgcaacctc
300
tgccctcccg gttcaagc
318

```

```

<210> 122
<211> 89
<212> PRT
<213> Homo sapiens

```

```

<400> 122
Xaa Met Gly Gly Pro Gly Thr Ala Leu Val Pro Leu Phe Phe Leu Gly
1              5              10              15
Lys Lys Leu Ser Pro Lys Gly Phe Ala Ala Leu Gln Glu Ser Phe Leu
20              25              30
Val Ser Leu Gly Leu Phe Leu Cys Cys Val Arg Arg Ser Arg Ser Ser
35              40              45
Ser Ser Arg Val Gly Leu Phe Ile Phe Ile Gln Arg Lys Phe Val Gly
50              55              60
Cys Gly Tyr Tyr Phe Leu Phe Phe Leu Phe Phe Phe Cys Leu Glu Thr
65              70              75              80
Glu Ser Xaa Ser Val Ala Arg Leu Glu
      85

```

```

<210> 123
<211> 338
<212> DNA
<213> Homo sapiens

```

```

<400> 123
acgcgtctag ggtagaaatc aactccagta actgtcatc aacctcagca atgctggggg
60

```

cgggcagagg cagggcagct gtgtgccaca ttctgtccag ggtgggtcag gccccggctc
 120
 tcacctctcc tcctccctgc ttggaacctg tggaacaaag ggccccctgca ccccaactca
 180
 ttctctcttg ccacataagg gcttcaagtc atgtgtctcc ctctgcttgg gttgcttttt
 240
 ctccctctgc ttgggtcact gttcacacca ctggccactt tcctcaggga agggccctca
 300
 ctgcccacac acctaaacat gccccctgct cctccata
 338

<210> 124

<211> 96

<212> PRT

<213> Homo sapiens

<400> 124

Met	Leu	Gly	Thr	Gly	Arg	Gly	Arg	Ala	Ala	Val	Cys	His	Ile	Pro	Ala
1			5					10					15		
Arg	Ala	Gly	Gln	Ala	Pro	Ala	Leu	Thr	Thr	Pro	Pro	Pro	Cys	Phe	Glu
			20					25					30		
Pro	Val	Glu	Gln	Arg	Ala	Pro	Ala	Pro	Gln	Leu	Ile	Pro	Leu	Cys	His
			35				40					45			
Ile	Arg	Ala	Ser	Ser	His	Ala	Val	Pro	Ser	Ala	Trp	Val	Ala	Phe	Ser
			50			55				60					
Pro	Ser	Ala	Trp	Val	Thr	Val	His	Thr	Thr	Gly	His	Phe	Pro	Gln	Gly
					70					75				80	
Arg	Ala	Leu	Thr	Ala	His	Thr	Pro	Lys	His	Ala	Pro	Cys	Ser	Ser	Ile
				85					90					95	

<210> 125

<211> 280

<212> DNA

<213> Homo sapiens

<400> 125

ccatggacct ggccagccac catcacctgc ctctgcctc acccaccctg ggtgcctgcc
 60
 ggcaaggatt ggagggcaga ctgctggagc gtgagaccag gccaatctgt cttttctggga
 120
 accttcagcc tccaactgga gctgactgtc aactttcggt tgagaagtca cttttctgca
 180
 ttcccaccac actatctatc tgtgcaatc ggcagcgtga cagcactcac cttattgagg
 240
 gcttctgtctg tcctggccca ttctggatag gctgatcta
 280

<210> 126

<211> 92

<212> PRT

<213> Homo sapiens

<400> 126

Met Asp Leu Ala Ser His His His Leu Pro Pro Ala Ser Pro Thr Leu

```

      1             5             10             15
Gly Ala Cys Arg Gln Gly Leu Glu Gly Arg Leu Leu Glu Arg Glu Thr
      20             25             30
Arg Pro Ile Cys Leu Ser Gly Asn Leu Gln Pro Pro Thr Gly Ala Asp
      35             40             45
Cys Gln Leu Ser Gly Glu Lys Ser Leu Phe Cys Ile Pro Thr Thr Leu
      50             55             60
Ser Ile Cys Ala Ile Arg Gln Arg Asp Ser Thr His Leu Ile Glu Gly
      65             70             75             80
Phe Cys Cys Pro Gly Pro Phe Trp Ile Gly Leu Ile
      85             90

```

<210> 127

<211> 444

<212> DNA

<213> Homo sapiens

<400> 127

```

cgctgatcgc ccgtggcgga gggccgcggc gcgactcga tcgccagct gacaaccgag
60
ctgcaaagcc gtcactgcc tcgggagcag atcacgtccg tcagcatcga catgtcgcca
120
gcgttcatca ggggctgcgc cgagcacctg cccaacgcgc gcgtcacctt cgacaagttc
180
cacgtcatcg ggcacgccaa tgcggccgtg gacaggatgc gccgcacga gcaagcgagc
240
gacaagtccc tcaaggggat gcgctggctg ctgctgaaga accgcgccag cctcaagccc
300
gaggctgcgc ccgatctgga tgccctgac gccaggatgg ccactgtgcy caccgcgcgc
360
gcctgggtct acaaggagca gctgcgcgag atcctcgccg gcaagcagat caacgtggca
420
cgcgacatgc tcaagcactg gtgc
444

```

<210> 128

<211> 148

<212> PRT

<213> Homo sapiens

<400> 128

```

Arg Val Ile Ala Val Ala Glu Gly Arg Gly Ala Asp Ser Ile Ala Gln
      1             5             10             15
Leu Thr Thr Glu Leu Gln Ser Arg His Cys Pro Ala Glu Gln Ile Thr
      20             25             30
Ser Val Ser Ile Asp Met Ser Pro Ala Phe Ile Arg Gly Cys Ala Glu
      35             40             45
His Leu Pro Asn Ala Arg Val Thr Phe Asp Lys Phe His Val Ile Gly
      50             55             60
His Ala Asn Ala Ala Val Asp Arg Met Arg Arg Ile Glu Gln Arg Ser
      65             70             75             80
Asp Lys Ser Leu Lys Gly Met Arg Trp Ser Leu Leu Lys Asn Arg Ala
      85             90             95
Ser Leu Lys Pro Glu Ala Ala Ala Asp Leu Asp Ala Leu Ile Ala Arg

```

```

          100          105          110
Met Ala Thr Val Arg Thr Ala Arg Ala Trp Val Tyr Lys Glu Gln Leu
      115          120          125
Arg Glu Ile Leu Ala Arg Lys Gln Ile Asn Val Ala Arg Asp Met Leu
      130          135          140
Lys His Trp Cys
145

```

```

<210> 129
<211> 291
<212> DNA
<213> Homo sapiens

```

```

<400> 129
gaggaggagac gtaccgtccc cgttatagcc aagctcgaga agccgcaagc tatcgagaac
60
ttggacgaga ttattgacgt ctttgacgcc gtcattggttg cccgtggcga tatggccgtc
120
gagtgcccg ctcgaggaagt tccgctgac caaaagcaga tcacgagaa ggctcgctta
180
caggctaagc ccgtcattgt gggcaccag atgcttgagt cgatgatcca cgctccccgt
240
cgcacccgcg ctgaggccgc cgacgtcgcg aacgccatcc ttgacggcgc g
291

```

```

<210> 130
<211> 97
<212> FRT
<213> Homo sapiens

```

```

<400> 130
Glu Glu Gly Arg Thr Val Pro Val Ile Ala Lys Leu Glu Lys Pro Gln
1      5      10      15
Ala Ile Glu Asn Leu Asp Glu Ile Ile Asp Val Phe Asp Ala Val Met
      20      25      30
Val Ala Arg Gly Asp Met Ala Val Glu Cys Pro Leu Glu Glu Val Pro
      35      40      45
Leu Ile Gln Lys Gln Ile Ile Glu Lys Ala Arg Leu Gln Ala Lys Pro
      50      55      60
Val Ile Val Ala Thr Gln Met Leu Glu Ser Met Ile His Ala Pro Arg
      65      70      75      80
Pro Thr Arg Ala Glu Ala Ala Asp Val Ala Asn Ala Ile Leu Asp Gly
      85      90      95
Ala

```

```

<210> 131
<211> 416
<212> DNA
<213> Homo sapiens

```

```

<400> 131
tcggagcgt ccgtggccct catgggtgtg tcagcgtggt tgctgtctcg ggccgcagag
60

```

attccaccgg tgctctacct ggaggccgca gccgtcgggg ttcgattctt cggcatctcc
 120
 cgcggtgtct tccgctacgc cgaacgtctg gtaggccacg acctggctct gcggatgcag
 180
 ggggcattgc gtagtgcgggt ctacgaccgg ctgtcacgta ccnaccctgc tggggnnactg
 240
 cgccgggggt acctgctgggt acgggttact gccgacgtcg acgggtgttt ggacatggtc
 300
 gtgcgggtga tcgttccggc gtgcgcgtca agcctcgtca tcattggcac cacggctcctt
 360
 cttgtccga gagaagggtg agttttctta gccggattcc aacacagcct gggggc
 416

<210> 132

<211> 126

<212> PRT

<213> Homo sapiens

<400> 132

Ser Gly Ala Ser Val Ala Leu Met Gly Val Ser Ala Trp Leu Leu Ser
 1 5 10 15
 Arg Ala Ala Glu Ile Pro Pro Val Leu Tyr Leu Glu Ala Ala Ala Val
 20 25 30
 Gly Val Arg Phe Phe Gly Ile Ser Arg Gly Val Phe Arg Tyr Ala Glu
 35 40 45
 Arg Leu Val Gly His Asp Leu Ala Leu Arg Met Gln Gly Ala Leu Arg
 50 55 60
 Met Arg Val Tyr Asp Arg Leu Ser Arg Thr Xaa Pro Ala Gly Xaa Arg
 65 70 75 80
 Arg Arg Gly Asp Leu Leu Val Arg Val Thr Ala Asp Val Asp Ala Val
 85 90 95
 Leu Asp Met Val Val Arg Val Ile Val Pro Ala Cys Ala Ser Ser Leu
 100 105 110
 Val Ile Ile Gly Thr Thr Val Leu Leu Cys Pro Arg Glu Gly
 115 120 125

<210> 133

<211> 327

<212> DNA

<213> Homo sapiens

<400> 133

gccgttgcta tcgctgctgg tatgctgca gacgtcactg tttttgatat caatatcgct
 60
 gcgttgaaga gactcgccga catctaccag ggtcgtgttc acacagtagt atccaccgcc
 120
 gccgaattg cgaaggcgct agaaaccgct gacgttgtga tcggttctgt ccttattccg
 180
 ggtagttcta ccccgaaagct tggtactacc gatattggtg ctcacatgca gcctgggtct
 240
 gttcttattg atattgctat agaccaaggc ggctgcttcg aggattcgca cccaccact
 300
 tacgatgacc ccactttcac tgtgcac
 327

<210> 134
 <211> 109
 <212> PRT
 <213> Homo sapiens

<400> 134
 Ala Val Ala Ile Ala Ala Gly Met Arg Ala Asp Val Thr Val Phe Asp
 1 5 10 15
 Ile Asn Ile Ala Ala Leu Lys Arg Leu Ala Asp Ile Tyr Gln Gly Arg
 20 25 30
 Val His Thr Val Val Ser Thr Arg Ala Glu Ile Ala Lys Ala Leu Glu
 35 40 45
 Thr Ala Asp Val Val Ile Gly Ser Val Leu Ile Pro Gly Ser Ser Thr
 50 55 60
 Pro Lys Leu Val Thr Thr Asp Met Val Ala His Met Gln Pro Gly Ser
 65 70 75 80
 Val Leu Ile Asp Ile Ala Ile Asp Gln Gly Gly Cys Phe Glu Asp Ser
 85 90 95
 His Pro Thr Thr Tyr Asp Asp Pro Thr Phe Thr Val His
 100 105

<210> 135
 <211> 560
 <212> DNA
 <213> Homo sapiens

<400> 135
 taagatgtgg tccctgccctg ttccctgaagg ggctgcagct ctgatggaaa atacagggat
 60
 ttacactcag ggctacagcc acggggggct gaggcccaag gctgcaatct cgggggaagg
 120
 ggaagtgtggc ttttctctgt ggattggaaa catcctcttg gaggcaaaga ctttctctgg
 180
 attctacaga cttcccggga ttttttagatt agaattattgg gggcaaagga ggctgtcttg
 240
 ttttaaagca atgctacata gacacagtgg ggaagacctg gttcgacggc agataagcag
 300
 tgggtgatgg gcttgaggag gagagtcagg gcaaagtcta agactgagca gaaaggaatt
 360
 cccccatctc ccatggataa gtacgtttcta gaacattctc tttgggtcta atactctgaa
 420
 atgacattct gtcttcacgc tcgagagaga attacttccac tggctccact tggagtgcga
 480
 gtgttcagac accaagcctg actgggaggg ttccgttttc ttaacacctt cccaccgccg
 540
 acttccaagt cccacgcgt
 560

<210> 136
 <211> 100
 <212> PRT
 <213> Homo sapiens

<400> 136

```

Met Trp Ser Cys Pro Val Pro Glu Gly Ala Ala Ala Leu Met Glu Asn
 1             5             10             15
Thr Gly Ile Tyr Thr Gln Gly Tyr Ser His Gly Gly Leu Arg Pro Lys
      20             25             30
Ala Ala Ile Ser Gly Glu Gly Glu Val Gly Phe Ser Trp Trp Ile Gly
      35             40             45
Asn Ile Leu Leu Glu Ala Lys Thr Phe Pro Gly Ser Tyr Arg Leu Pro
      50             55             60
Gly Ile Phe Arg Leu Glu Tyr Trp Gly Gln Arg Arg Leu Ser Cys Phe
65             70             75             80
Lys Ala Met Leu His Arg His Ser Gly Glu Asp Leu Val Arg Arg Gln
      85             90             95

Ile Ser Ser Gly
      100

```

<210> 137

<211> 429

<212> DNA

<213> Homo sapiens

<400> 137

```

accgggtgga tggcctgcag gccaaagcgt tctcgcaaac tcagcaggcc ttcagcgcaa
60
gaggcaaaaca gctggtgcgc cacctgcttg aggtccaccg attgcgcatt gcccttgagg
120
aaggcgcgcgc agttgggttt gtccgccact tggctgcgga acaggtcttc gacaaaaccg
180
gactgctggc gggtcgcaac gcgcatgac gccagcgccct ggctggcgcc ctggtcgagc
240
cagcgcgtcg gcagttgggt ggcccgggtg ataccgacct tgatccccga cgaattggcc
300
aggtacacca catggtcggt catgcagaat gtttcgcccc agcggggatc acggcaagtg
360
ccggcgctcgt aatggcaacg ttccggggctc atgatgcaca ggtcacactg ggccagcttg
420
gtcatgcccc
429

```

<210> 138

<211> 141

<212> PRT

<213> Homo sapiens

<400> 138

```

Met Thr Lys Leu Ala Gln Cys Asp Leu Cys Ile Met Ser Pro Glu Arg
 1             5             10             15
Cys His Tyr Asp Ala Gly Thr Cys Arg Asp Pro Gly Trp Gly Glu Thr
      20             25             30
Phe Cys Met Thr Asp His Val Val Tyr Leu Ala Asn Ser Ser Gly Ile
      35             40             45
Lys Val Gly Ile Thr Arg Ala Thr Gln Leu Pro Thr Arg Trp Leu Asp
      50             55             60
Gln Gly Ala Ser Gln Ala Leu Pro Ile Met Arg Val Ala Thr Arg Gln

```



```

65              70              75              80
Gln Ser Gly Phe Val Glu Asp Leu Phe Arg Ser Gln Val Ala Asp Lys
              85              90              95
Thr Asn Trp Arg Ala Leu Leu Lys Gly Asp Ala Gln Ser Val Asp Leu
              100              105              110
Lys Gln Val Arg Asp Gln Leu Phe Ala Ser Cys Ala Glu Gly Leu Leu
              115              120              125
Ser Leu Gln Glu Arg Phe Gly Leu Gln Ala Ile Gln Pro
              130              135              140

```

<210> 139

<211> 341

<212> DNA

<213> Homo sapiens

<400> 139

```

acgcgtcggtt tgaaggcttg atccgcacgt ccaattcgct ttgcgccaat gcgccgcagc
60
ttgtgaacag cagaatcaag ccgctggtaa atcttctctg gagcttcata ggcgggggatg
120
ctacacgagc tggggagaca ctttgaaccc ggaattgtct gaataattct gtctcaaac
180
tttgcagcct gtaacgactg aggggttcgga tggaaaaaca catgctccag gatgggacgg
240
acggccactt caccgatctc ttcatagcc ttgctgttgt agaaatccag gtacgcgcgaa
300
tcgccagcgt cgagcacgac gcctgatgag tgcgggtcat t
341

```

<210> 140

<211> 113

<212> PRT

<213> Homo sapiens

<400> 140

```

Met Thr Arg Thr His Gln Ala Ser Cys Ser Thr Leu Ala Ile Arg Ala
1              5              10              15
Thr Trp Ile Ser Thr Asn Ala Arg Ala Met Lys Arg Ser Val Lys Trp
              20              25              30
Pro Ser Val Pro Ser Trp Ser Met Cys Phe Ser Ile Arg Thr Leu Ser
              35              40              45
Arg Tyr Arg Leu Gln Arg Phe Glu Thr Glu Leu Phe Arg Gln Phe Arg
              50              55              60
Val Gln Ser Val Ser Pro Ala Arg Val Ala Ser Pro Pro Met Lys Leu
              65              70              75              80
Pro Gly Arg Phe Thr Ser Gly Leu Ile Leu Leu Phe Thr Ser Cys Gly
              85              90              95
Ala Leu Ala Gln Ser Glu Leu Asp Val Arg Ile Lys Pro Ser Asn Asp
              100              105              110
Ala

```

<210> 141

<211> 324

<212> DNA

<213> Homo sapiens

<400> 141

gaattcctctt tggatagctt cgggtaaatg ggtacagcaa atatcaggag cgcaaccgca
 60
 acccttactt actggtacat gaacaccatt tacattacag ctatcgtact caccaccggt
 120
 catgtgaaca gacacataac tgaaggttt ataaaccaca gtctcacggt acgtatgacc
 180
 gtcaactgtg aacaccgcta agtaaatagcc tgcgggggct tgcatgaact cctttgacca
 240
 tgcgtaataa atacgtccgt cattagtcac acctgatggg gcgaaacaaa aagaacggca
 300
 gcagttatca ccgcccatac gcgt
 324

<210> 142

<211> 106

<212> PRT

<213> Homo sapiens

<400> 142

Met	Gly	Gly	Asp	Asn	Cys	Cys	Arg	Ser	Phe	Cys	Phe	Ala	Pro	Ser	Gly
1				5				10					15		
Val	Thr	Asn	Asp	Gly	Arg	Ile	Tyr	Tyr	Ala	Trp	Ser	Lys	Glu	Phe	Met
		20					25					30			
Gln	Ala	Pro	Ala	Gly	Tyr	Tyr	Leu	Ala	Val	Phe	Thr	Val	Asp	Gly	His
		35				40					45				
Thr	Tyr	Arg	Glu	Thr	Val	Val	Tyr	Lys	Pro	Phe	Ser	Tyr	Val	Ser	Val
		50			55					60					
His	Met	Thr	Trp	Gly	Glu	Tyr	Asp	Ser	Cys	Asn	Val	Asn	Gly	Val	His
65				70				75					80		
Val	Pro	Val	Ser	Lys	Gly	Cys	Gly	Cys	Ala	Pro	Asp	Ile	Cys	Cys	Thr
			85					90					95		
His	Leu	Pro	Glu	Ala	Ile	Gln	Glu	Glu	Phe						
		100					105								

<210> 143

<211> 1325

<212> DNA

<213> Homo sapiens

<400> 143

nacgctgga tctgccagct gagcctggag ctgtgcagge agctgccctg ctacgatgag
 60
 gcacccagg agaagaactt cctgtacaaa tgcataggca ccaccctggg tgctgcttca
 120
 agtaaggagg tggtaggaa gcaccttcaa gagctgctgg agacggccag ataccaggag
 180
 gaggcagaag gcgagggcct cgctctgtgc ttgggatctc gtgccatctc ccacctcgag
 240
 gacacgctgg cccagctgga ggacttcgtg aggtcagagg tcttcagaaa atccattggg
 300

attctcaaca tttttaagga tcgaagtgag aacgaagtgg agaaggtgaa gagtgtctgtg
 360
 atccctgtgct atgggcacgt ggcggcccggt gcccccgggt agctggtgct ggccaaaggtg
 420
 gagtgcagaa tcctccggaa catcntgccg gcacttcagc acnncaagga cccagccctg
 480
 aagctgtgtgc ttgtccagag tgtgtgcatg gtcagcccggt ccatctgcag cagcaccacg
 540
 gctggctcct tccacttcac ccggaagca gagtgtgtgg cacagatgat ggagttcatc
 600
 agggcagagc ccccggaact cttgaggaca cctattcgga agaaagccat gctcacctgc
 660
 acttacttgg tctccgtgga gccagcgctg gacgagcagg ccgggcgga tgtgatccat
 720
 ggctgcctgc acagcatcat ggcctgtgtg cctgagccca aggaggagga cggaggctgc
 780
 cagaagtcct tgtatctgga gacactgcac gcccttgagg atctgtgtac gagcctcctg
 840
 cagcggaaca tgacccccca aggcctgcag atcatgattg agcacctgag cccatggatc
 900
 aagtcaccaa gaggtcacgt agcggcgctg gccctaggcc tgagcgccct cctcgtgcgc
 960
 tacttcctgg agcacctgcg tgtcagtggt gcccaagtag ataccaggtt tccatctgag
 1020
 cccaggtatc ttgtcaatgg cctcgtgtgc cttccacaac ctggggccttc tcatcgccct
 1080
 cttctcccca cgtgtgtcgg acctgtggcc tgccaccgcg caggaggccg tggactgtgt
 1140
 ctactccctg ctgtacctcc agctcggcta tgagggtctc tcccgggact acccgcatga
 1200
 cgtggcggag cggctcctca gcctcaagga cggcctcgtg caccctgacc ccgccattct
 1260
 cttccacacc tgccacagtg taggccagat tattgccaaag cgcctccccc cagcccttca
 1320
 cgcgt
 1325

<210> 144

<211> 390

<212> PRT

<213> Homo sapiens

<400> 144

Xaa Ala Trp Ile Cys Gln Leu Ser Leu Glu Leu Cys Arg Gln Leu Pro
 1 5 10 15
 Cys Tyr Asp Glu Ala Pro Gln Glu Lys Asn Phe Leu Tyr Lys Cys Ile
 20 25 30
 Gly Thr Thr Leu Gly Ala Ala Ser Ser Lys Glu Val Val Arg Lys His
 35 40 45
 Leu Gln Glu Leu Leu Glu Thr Ala Arg Tyr Gln Glu Glu Ala Glu Arg
 50 55 60
 Glu Gly Leu Ala Cys Cys Phe Gly Ile Cys Ala Ile Ser His Leu Glu
 65 70 75 80
 Asp Thr Leu Ala Gln Leu Glu Asp Phe Val Arg Ser Glu Val Phe Arg

85 90 95
 Lys Ser Ile Gly Ile Leu Asn Ile Phe Lys Asp Arg Ser Glu Asn Glu
 100 105 110
 Val Glu Lys Val Lys Ser Ala Leu Ile Leu Cys Tyr Gly His Val Ala
 115 120 125
 Ala Arg Ala Pro Arg Glu Leu Val Leu Ala Lys Val Glu Ser Asp Ile
 130 135 140
 Leu Arg Asn Ile Xaa Pro Ala Leu Gln His Xaa Lys Asp Pro Ala Leu
 145 150 155 160
 Lys Leu Cys Leu Val Gln Ser Val Cys Met Val Ser Arg Ala Ile Cys
 165 170 175
 Ser Ser Thr Gln Ala Gly Ser Phe His Phe Thr Arg Lys Ala Glu Leu
 180 185 190
 Val Ala Gln Met Met Glu Phe Ile Arg Ala Glu Pro Pro Asp Ser Leu
 195 200 205
 Arg Thr Pro Ile Arg Lys Lys Ala Met Leu Thr Cys Thr Tyr Leu Val
 210 215 220
 Ser Val Glu Pro Ala Leu Asp Glu Gln Ala Arg Ala Asp Val Ile His
 225 230 235 240
 Gly Cys Leu His Ser Ile Met Ala Leu Leu Pro Glu Pro Lys Glu Glu
 245 250 255
 Asp Gly Gly Cys Gln Lys Ser Leu Tyr Leu Glu Thr Leu His Ala Leu
 260 265 270
 Glu Asp Leu Leu Thr Ser Leu Leu Gln Arg Asn Met Thr Pro Gln Gly
 275 280 285
 Leu Gln Ile Met Ile Glu His Leu Ser Pro Trp Ile Lys Ser Pro Arg
 290 295 300
 Gly His Val Ala Ala Arg Ala Leu Gly Leu Ser Ala Leu Leu Val Arg
 305 310 315 320
 Tyr Phe Leu Glu His Leu Arg Val Ser Gly Ala Gln Val Asp Thr Arg
 325 330 335
 Phe Pro Ser Glu Pro Arg Ile Leu Cys Asn Gly Pro Gly Ala Leu Pro
 340 345 350
 Gln Pro Gly Pro Ser His Arg Pro Leu Leu Pro Thr Val Cys Gly Pro
 355 360 365
 Val Ala Cys His Pro Pro Gly Gly Arg Gly Leu Cys Leu Leu Pro Ala
 370 375 380
 Val Pro Pro Ala Arg Leu
 385 390

<210> 145

<211> 802

<212> DNA

<213> Homo sapiens

<400> 145

cgcgcgtctc ggtccggctc agtgcgctgt tgcctgcgct agaacacgag gctgcgcaag
 60
 cataagcaga cgtagagagt ggtcacatcc atgtcgatgg tgtgcgcgta atgaaggtct
 120
 acatcacccct ggtgaaggcc tgcaccacta gcgtcggcac catttccccg cgtcggacaa
 180
 gacatcatgc cccatattctt gacagaatgt ctgacatgag tatgccacgc cgagcagcac
 240

cagaggacga caccgatctg gcggacgccg cccgttcatg gcgcagatac ctcatcctcg
 300
 tcatttgtgg cgttatcgtc gctgtcctcg gactaggcat ttctgggtat cttgcgtggt
 360
 ggctcattgtg cgatcaagct gccgggggtct gtcagcgtgg tgaacccgtt atgtactggt
 420
 gttccgtggt ctctctggcc attctcggac tcattatcgg ggtcttgacg cagatctggc
 480
 tggagaagcg ctggtggcac atgcttgcca tcgtcatccc ggctgttttc atcgtcgcgg
 540
 gtatcttttt ttggctcgcc gtctaagaag gggcgctcac gattccacaa acgacacagg
 600
 tattgatctc cgttttatcg gctcctagca gccgtggtea acgtatcgct atcaagcgat
 660
 acaggactcg tcgttcgcat cgttgttggg ctgctgggaa acaatcccg cgatctactc
 720
 ggctaccgcc agacagttca ctcaaaccc ctacgcggcg cgacagatc aaatccatt
 780
 ctcatagac ggcccacacc ac
 802

<210> 146
 <211> 151
 <212> PRT
 <213> Homo sapiens

<400> 146
 Met Lys Val Tyr Ile Thr Leu Val Lys Ala Cys Thr Thr Ser Val Gly
 1 5 10 15
 Thr Ile Ser Pro Arg Arg Thr Arg His His Ala Pro Tyr Leu Asp Arg
 20 25 30
 Met Ser Asp Met Ser Met Pro Arg Arg Ala Ala Pro Glu Asp Asp Thr
 35 40 45
 Asp Leu Ala Asp Ala Ala Arg Ser Trp Arg Arg Tyr Leu Ile Leu Val
 50 55 60
 Ile Cys Gly Val Ile Val Ala Val Leu Gly Leu Gly Ile Phe Gly Tyr
 65 70 75
 Leu Ala Trp Trp Ser Leu Cys Asp Gln Ala Ala Gly Val Cys Gln Arg
 85 90 95
 Gly Glu Pro Val Met Tyr Trp Cys Ser Val Val Ser Leu Ala Ile Leu
 100 105 110
 Gly Leu Ile Ile Gly Val Leu Thr Gln Ile Trp Leu Glu Lys Arg Trp
 115 120 125
 Trp His Met Leu Ala Ile Val Ile Pro Ala Val Phe Ile Val Ala Gly
 130 135 140
 Ile Phe Phe Trp Leu Ala Val
 145 150

<210> 147
 <211> 368
 <212> DNA
 <213> Homo sapiens

<400> 147

acgcgtgaaa acggtatgac tcttctggcc ttagtagatc tgtctaaaaa acccgatgag
 60
 tttacacagt gggcattagt agcccgcat gtctcatgaca ttcttggtct acgaaaagtt
 120
 attggtcaga aagtaaccttg tgttgcagtg acggggtcgg aaaaggtgct tcataaaaaag
 180
 gattactggg atctagcaac acctatgccat attgcgtggg gtacaacgga ccgaacagtt
 240
 attgctgatg cagcagctac aatccccacc acggagtggg atatccttgc aagactacgt
 300
 ccacgcctag aagagggttcg caagcaacgt aatgatgtat tgctcctcaa cgaggaggat
 360
 ccccccta
 368

<210> 148

<211> 117

<212> PRT

<213> Homo sapiens

<400> 148

Met	Thr	Leu	Leu	Ala	Leu	Val	Asp	Leu	Ser	Lys	Lys	Pro	Asp	Glu	Phe
1				5					10					15	
Thr	Gln	Trp	Ala	Leu	Val	Ala	Arg	Asp	Val	His	Asp	Ile	Pro	Gly	Leu
			20				25						30		
Arg	Lys	Val	Ile	Gly	Gln	Lys	Val	Pro	Cys	Val	Ala	Val	Thr	Gly	Ser
		35				40						45			
Glu	Lys	Val	Leu	His	Lys	Lys	Asp	Tyr	Trp	Asp	Leu	Ala	Thr	Pro	Met
		50				55					60				
Pro	Ile	Ala	Trp	Gly	Thr	Thr	Asp	Arg	Thr	Val	Ile	Ala	Asp	Ala	Arg
		65			70					75				80	
Arg	Thr	Ile	Pro	Thr	Thr	Glu	Trp	Asp	Ile	Leu	Ala	Arg	Leu	Arg	Pro
				85				90						95	
Arg	Leu	Glu	Glu	Val	Arg	Lys	Gln	Arg	Asn	Asp	Val	Leu	Leu	Leu	Asn
				100				105						110	
Glu	Glu	Asp	Pro	Pro											
				115											

<210> 149

<211> 407

<212> DNA

<213> Homo sapiens

<400> 149

nngctagcat ggaccctagt cacacaggca gccatacccg aggtcaaaagt gaccattttt
 60
 cctaataatgg ccgctcagat ccaatacttt gaagattcgt ccgtggttat atggcacgat
 120
 gcggtgatg gtatcgtgta ccgaagtgcg gatgaaggca agtcgtgggc cccaattaag
 180
 gggcctgaac agggtcaggc gcaccttttc gtgtccatc cctacgacaa gactcaagcg
 240
 tatattctga cgcgcagcac tcagcattgg cgcacgtcga accgtggcga gacgtggcga
 300

tcattctcaa cgctcatcc gctacgacc ttgaaagcta tgcctctgga ctttcatccg
 360
 acgcatcatg actggatcct ttacacgggc caggcttgca cggtaaa
 407

<210> 150
 <211> 135
 <212> PRT
 <213> Homo sapiens

<400> 150
 Xaa Leu Ala Trp Thr Leu Val Thr Gln Ala Ala Ile Pro Glu Val Lys
 1 5 10 15
 Val Thr His Phe Pro Asn Met Ala Ala Gln Ile Gln Tyr Phe Glu Asp
 20 25 30
 Ser Ser Val Val Ile Trp His Asp Ala Val Asp Gly Ile Val Tyr Arg
 35 40 45
 Ser Ala Asp Glu Gly Lys Ser Trp Ala Pro Ile Lys Gly Pro Glu Gln
 50 55 60
 Gly Gln Ala His Leu Phe Val Leu His Pro Tyr Asp Lys Thr Gln Ala
 65 70 75 80
 Tyr Ile Leu Thr Arg Ser Thr Gln His Trp Arg Thr Ser Asn Arg Gly
 85 90 95
 Glu Thr Trp Gln Ser Phe Ser Thr Pro Pro His Pro Pro Thr Thr Leu Lys
 100 105 110
 Ala Met Pro Leu Asp Phe His Pro Thr His His Asp Trp Ile Leu Phe
 115 120 125
 Thr Gly Gln Ala Cys Thr Val
 130 135

<210> 151
 <211> 448
 <212> DNA
 <213> Homo sapiens

<400> 151
 accggtgtcc gtggctattg ccccgaaatgg tccccatccg cgccccggg aactccctcg
 60
 gcttttcgcg catccaggtc ccagcggcca gctactgggt cgccccgagc ccctaggtgc
 120
 cagagcgggtg gtcggcgggg ctctgcccc gctcgggtc ctccctcctc cccaccagaa
 180
 ggaaaaactt gggcccttcg agaacctgt ggaatgttct ttgtaataca ctgtacatcc
 240
 gctttcacgg cagcgctcgt tgcaaaatcg cgggttttcg ggccttgagg caaattgcgc
 300
 ttgtcagcgg cgacgtcagg aggacaaggg gaggggttcg cggctgaaac tgcagcttcg
 360
 cagcacagag ccatttttagg ctgctcccca cctcgcgggg cccatgggaa gccggccccg
 420
 ggaggcgcg gctgcattga tattcgac
 448

<210> 152

<211> 149

<212> PRT

<213> Homo sapiens

<400> 152

```

Thr Gly Val Arg Gly Tyr Cys Pro Glu Trp Ser Pro Ser Ala Ser Pro
 1           5           10           15
Gly Thr Pro Ser Ala Phe Arg Ala Ser Arg Ser Pro Ala Pro Ala Thr
 20           25           30
Gly Ala Pro Arg Ala Pro Arg Cys Gln Ser Gly Gly Arg Pro Gly Ser
 35           40           45
Cys Pro Val Ser Ala Pro Pro Ser Ser Pro Pro Glu Gly Lys Thr Trp
 50           55           60
Ala Leu Arg Glu Pro Cys Gly Met Phe Phe Val Ile Asn Cys Thr Ser
 65           70           75           80
Ala Ser Thr Ala Arg Pro Arg Ala Lys Ser Arg Val Ser Gly Pro Trp
 85           90           95
Ser Lys Leu Arg Leu Ser Ala Ala Thr Ser Gly Gly Gln Gly Glu Gly
100           105           110
Phe Ala Ala Glu Thr Ala Ala Ser Gln His Arg Ala Ile Leu Gly Cys
115           120           125
Ser Pro Pro Arg Gly Ala His Gly Lys Pro Ala Pro Gly Gly Arg Gly
130           135           140
Cys Met Asp Ile Arg
145

```

<210> 153

<211> 440

<212> DNA

<213> Homo sapiens

<400> 153

```

nnntgggtcc atgtatgtgt gtgtatatga gggagacacg cagggtgtgtg tccgagtggtg
60
tgtccatggg tccatgtatg tgtgtgtata tgtgggggaa cagggtgtgtg tccgagtggtg
120
tgcattgggtc cgtgtatatg cgtgtatatata tgcggggata tgtatatgtg tgtgtgtatg
180
aacagggtgta agtggggagc actcagggtg gtctgtgtgtg gttcgtgtac acgtgtgttaa
240
gtgggtgtgac atgaaggggt gtgtgtgtgtc gtgtgtagtg ttgcgtgcat gcacacatgc
300
atgtgtgttac tggggcatcc aagcccctgg tctccactcc attccacctc acgcctaccc
360
ccttgatctc tgcgccagc cttggctgtg ctcccctgct gtatgcacgt ggggtgtctgc
420
acgtgggtgt ctgcacgcgt
440

```

<210> 154

<211> 69

<212> PRT

<213> Homo sapiens

<400> 154
 Gly Arg His Ala Gly Val Cys Pro Ser Val Cys Pro Trp Val His Val
 1 5 10 15
 Cys Val Cys Ile Cys Gly Gly Thr Gly Val Cys Pro Ser Val Cys Met
 20 25 30
 Gly Pro Cys Ile Cys Val Tyr Ile Cys Gly Asp Met Tyr Met Cys Val
 35 40 45
 Cys Met Asn Arg Cys Lys Trp Gly Ala Leu Arg Cys Val Cys Val Cys
 50 55 60
 Ser Cys Thr Arg Val
 65

<210> 155
 <211> 344
 <212> DNA
 <213> Homo sapiens

<400> 155
 acgcgtatcgc accaccatgt cgctcgtcacc acggcaagcg ctctcggcgg gcgagaacga
 60
 gtgaacatcgg ccgagttgat ggccgatgcc gcgaccggca cgaaacctgc ctacctacag
 120
 cgatcttctc cctcgatcac ctcgcttgaa gtggacaggg aacaaagaca ctgagacaac
 180
 ggcgcgcagg aagtaaaaag ttcgctctcc gatcacggcc gtcgcgcgag tgcacaggga
 240
 gaactggggc cctcgcaagc tacgccacgc cgatccatgc ccccgccctg atcttcgcgc
 300
 tcctctacct ccccttacc gatcagcatt atatccgacg taga
 344

<210> 156
 <211> 92
 <212> FRT
 <213> Homo sapiens

<400> 156
 Met Ala Glu Leu Met Ala Asp Ala Ala Thr Gly Thr Lys Pro Ser Tyr
 1 5 10 15
 Leu Gln Arg Ser Ser Ser Ser Ile Thr Ser Phe Glu Val Asp Arg Glu
 20 25 30
 Gln Arg His Ser Asp Asn Ala Pro Gln Glu Val Lys Ser Ser Leu Ser
 35 40 45
 Asp His Gly Arg Arg Ala Ser Ala Gln Gly Glu Leu Gly Thr Ser Gln
 50 55 60
 Ala Thr Pro Pro Arg Ser Met Pro Pro Val Ser Ser Ala Ser Ser
 65 70 75 80
 Thr Ser Pro Leu Pro Ile Ser Ile Ile Ser Asp Leu
 85 90

<210> 157
 <211> 6816
 <212> DNA
 <213> Homo sapiens

<400> 157
nnagatctcc aaagaaccaa aaggatcaat atttctggat tcctgtatgg gtgtcgttca
60
gaacaacaaa gtcaggcgtt ttgcttttga gctcaagatg caggacaaaa gtatgtatct
120
cttggcagca gacagtgaag tggaaatgga agaattggatc acaattctaa ataagatcct
180
ccagctcaac tttgaagctg caatgcaaga aaagcgaaat ggcgactctc acgaagatga
240
tgaacaaga aaattggaag gttctggttc cggtttagat agctacctgc cggaacttgc
300
caagagtgcga agagaagcag aaatcaaact gaaaagtga agcagagtca aactttttta
360
tttggaccga gatgcccaga agcttgactt ctcatcagct gagccagaag tgaagtcatt
420
tgaagagaag tttggaaaaa ggatccttgt caagtgcatt gatttatctt tcaatttgca
480
atgctgtgtt gccgaaaatg aagaaggacc cactacaaat gttgaacctt tctttgttac
540
tctatccctg ttgacataa aatacaacgc gaagatttct gccgatttcc acgtagacct
600
gaaccatttc tcagtggagg aaatgatcgc caccagctcc ccggcgctga tgaatggcag
660
tggggcgaaa cccaatctgc cctcaggggc atccttcagc aagccgccat gcagtatccg
720
aagcaggga tattttcagt cacttgtcct catccagata tatttcttgt ggccagaatt
780
gaaaaagtcc ttcaggggag catcacacat tgcgctgagc catatatgaa aagttcagac
840
tcttctaagg tggcccagaa ggtgctgaag aatgccaaag aggcattgcca aagactagga
900
cagtatagaa tgccatttgc ttgggcagca aggacattgt ttaaggatgc atctggaaat
960
cttgacaaaa atgccagatt ttctgccatc tacaggcaag acagcaataa gctatccaat
1020
gatgacatgc tcaagttact tgcagacttt cggaacctg agaagatggc taagctccca
1080
gtgatttttg gcaatctaga cattacaatt gataatgttt cctcagactt ccctaattat
1140
gttaattcat catacattcc caaaaacaa tttgaacct gcagtaaaac tcccatcagc
1200
tttgaagtgg aggaatttgt gccctgcata ccaaaacaca ctacgccta caccatctac
1260
accaatcacc tttacgttta tcctaagtac ttgaaatagc acagtcagaa gtcttttgcc
1320
aaggctagaa atattgcgat ttgcattgaa ttcaaagatt cagatgagga agactctcag
1380
ccccttaagt gcatttatgg cagacctggt gggccagttt tcacaagaag cgcctttgct
1440
gcagttttac accatcacca aaaccagaa ttttatgatg agattaaaa agagttgcc
1500
actcagctgc atgaaaagca ccacctgttg ctacattctt tccatgtcag ctgtgacaa
1560

tcaagtaaag gaagcacgaa gaagagggat gtcgttgaaa cccaagtggg ctactcctgg
1620
cttcccctcc tgaagacgag aagggtgggtg acaagcgagc agcacatccc ggctctggcg
1680
aaccttctct cgggctatct tggctaccag gagcttgga tgggcaggca ttatggctcg
1740
gaaattaaat gggtagatgg aggcaagcca ctgctgaaaa ttccactca tctggtttct
1800
acagtgtata ctcaggatca gcattttacat aattttttcc agtactgtca gaaaaccgaa
1860
tctggagccc aagccttagg aaacgaactt gtaagtacc ttaagagtct gcatgcgatg
1920
gaaggccacg tgatgatcgc cttcttgccc actatcctaa accagctgtt ccgagtcttc
1980
accagagcca cacaggaaga agtcgcggtt aacgtgactc gggtcattat tcatgtgggt
2040
gcccagtgcc atgaggaagg attggagagc cacttgaggt catatgttaa gtacgcgtat
2100
aaggctgagc catatgttgc ctctgaatac aagacagtgc atgaagaact gaccaaatcc
2160
atgaccacga ttctcaagcc ttctgctgat ttctcaccga gcaacaaact actgaagtac
2220
tcattgggttt tctttgatgt actgatcaaa tctatggctc agcatttgat agagaactcc
2280
aaagttaagt tgctgcaaaa ccagagattt cctgcattct atcatcatgc agtggaacc
2340
gttgtaaata tgctgatgcc acacatcact cagaagtttc gagataatcc agaggcatct
2400
aagaacgcga atcatagcct tgctgtcttc atcaagagat gtttcacctc catggacagg
2460
ggctttgtct tcaagcagat caacaactac attagctgtt ttgctcctgg agacccaaag
2520
accctctttg aatacaagtt tgaatttctc cgtgtagtgt gcaaccatga acattatatt
2580
ccgttgaaact tacciaatgcc atttgaaaaa ggcaggattc aaagatacca agacctccag
2640
cttgactact cattaacaga tgagtctcgc agaaaccact tcttggtgagg actgttactg
2700
agggagggtg ggacagccct ccaggagtgc cgggagggtc gtctgatcgc catcagtgtg
2760
ctcaagaacc tgctgataaa gcattctttt gatgacagat atgcttcaag gagccatgag
2820
gcaaggatag ccaccctcta cctgcctctg tttggtctgc tgattgaaaa cgtccagcgg
2880
atcaatgtga gggatgtgtc acccttccct gtgaacgcgg gcattgactgt gaaggatgaa
2940
tccctggctc taccagctgt gaatccgctg gtgacgcggc agaagggaaag caccctggac
3000
aacagccctg acaaggacct gctgggcgccc atctccggca ttgcttctcc atatacaacc
3060
tcaactccaa acatcaacag tgtgagaaat gctgattcga gaggatctct cataagcaca
3120
gattcgggta acagccttcc agaaaaggaat agtgagaaga gcaattccct ggataagcac
3180

caacaaagta gcacattggg aaattccgtg gttcgtctg ataaacttga ccagttctgag
3240
attaagagcc tactgatgtg ttctctctac atcttaaaga gcatgtctga tgatgctttg
3300
tttataattt ggaacaaggc ttcaacatct gaacttatgg atttttttac aatatctgaa
3360
gtctgcctgc accagttcca gtacatgggg aagcgataca tagccagaac aggaatgatg
3420
catgccagat tgcagcagct gggcagcctg gataactctc tcaacttttaa ccacagctat
3480
ggccactcgg acgcagatgt tctgcaccag tcattacttg aagccaacat tgctactgag
3540
gtttgcctga cagctctgga cagcctttct ctatttacct tggcggttaa gaaccagctc
3600
ctggccgacc atggacataa tcctctctatg aaaaaagttt ttgatgtcta cctgtgtttt
3660
cttcaaaaa atcagttctga aacggcttta aaaaatgtct tcaactgcctt aaggctctta
3720
atttataagt ttccctcaac attctatgaa gggagagcgg acatgtgtgc ggctctgtgt
3780
tacgagatcc tcaagtgtctg taactccaag ctgagctcca tcaggacgga ggctcccg
3840
ctgctctact tcctgatgag gaacaacttt gattacactg gaaagaagtc ctttgtccgg
3900
acacatttgc aagtcattcat atctgtcagc cagctgatag cagacgttgt tggcattggg
3960
ggaaccagat tccagcagtc cctgtccatc atcaacaact gtgccaacag tgaccggctt
4020
attaagcaca ccagcttctc ctctgatgtg aaggacttaa ccaaaaggat acgcacggtg
4080
ctaattggcca ccgccagat gaaggagcat gagaacgacc cagagatgct ggtggacctc
4140
cagttacagc tggccaaatc ctatgccagc acgcccagc tcaggaaagc gtggctcgac
4200
agcatggcca ggatccatgt caaaaatggc gatctctcag aggcagcaat gtgctatgtc
4260
cacgtaacag ccctagtggc agaatactc acacggaaag aagcagttca gtgggagccg
4320
ccccttctcc ccacagcca tagcgcctgc ctgaggagga gccggggagg cgtgtttaga
4380
caaggatgca ccgccttcag ggtcattacc ccaaacatcg acgaggagcg ctccatgatg
4440
gaagacgtgg ggaatgcagga tgtccatttc aacgaggatg tgctgatgga gtccttgag
4500
cagtgcgagc atggactctg gaaagccgag cgctacgagc tcattgccga catctacaaa
4560
cttatcatcc ccatttatga gaagcggagg gattttgaga ggctggccca tctgtatgac
4620
acgctgcacc gggcctacag caaagtgcac gaggtcatgc actcggggcg caggctcttg
4680
gggacctact tccgggtagc cttcttcggg caggcagcgc aataccagtt tacagacagt
4740
gaaacagatg tggagggatt ctttgaagat gaagatggaa aggagtatat ttacaaggaa
4800

cccaaactca caccgctgtc ggaaatttct cagagactcc ttaactgta ctcgataaa
4860
tttggttctg aaaatgtcaa aatgatacag gattctgcca aggtcaacct taaggatctg
4920
gattctaatg atgcctacat ccaggtgact cacgtcatcc ccttctttga cgaaaaagag
4980
ttgcaagaaa ggaaaacaga gtttgagaga tcccacaaca tccgcgcgtt catgtttgag
5040
atgcccattta cgcagaccgg gaagaggcag ggcgggggtg aagagcagtg caaacggcgc
5100
accatcctga cagccatata ctgcttccct tatgtgaaga agcgcacccc tgcattgtac
5160
cagcaccaca ctgacctgaa ccccatcgag gtggccattg acgagatgag taagaagggtg
5220
gcggagctcc ggcagctgtg ctctctggcc gaggtggaca tgatcaaaact gcagctcaaa
5280
ctccagggga gcgtgagtgt tcaggtcaat gctggccacc tagcatatgc gcgagctttc
5340
ttagatgata caaacacaaa gcgatatcct gacaataaag tgaagctgtg taaggaaagt
5400
ttcaggcaat ttgtggaagc ttgcggtcaa gccttagcgg taaacgaacg tctgattaaa
5460
gaagaccagg tcgagtatca ggaagaaatg aaagccaact acaggggaaat ggcgaggag
5520
ctttctgaaa tcatgcatga gcagatctgc cccttgagg agaagacgag cgtcttaccg
5580
aatcccttc acatcttcaa cgccatcagt gggactccaa caagcacaat ggttcacggg
5640
atgaccagct cgtcttcggt cgtgtgatta catctcatgg ccggtgtgtg gggacttgct
5700
ttgtcatttg caaactcagg atgctttcca aagccaatca ctggggagac cgagcacagg
5760
gaggaccaag ggaaggggga gagaaggaa ataaagaaca acgttatttc ttaacagact
5820
tcttatagga gttgtaagaa ggtgcacata tttttttaa tctcactggc aatattcaaa
5880
gttttctattg tgtcttaaca aaggtgtggt agacactctt gagctggagt tagattttat
5940
tcttcttgc agagttagtg tagaatagat ggcctacaga aaaaaaagg tctgggatct
6000
acatggcagg gagggtctga ctgacattga tgccctgggg acccttttgcc tcgaggctga
6060
gctggaaaat ctgaaaata tttttttttt cctgtggcac attcaggttg aatacaagaa
6120
ctatttttgt gactagtttt tgatgacctt agggaaactga ccattgtaat tttgtacca
6180
gtgaaccagg agatttagtg cttttatatt catttccttg catttaagaa aatatgaag
6240
cttaaggaa tatgtgagct taaaactagt caagcagttt agaaccaaa gcctatatta
6300
ataaccgcaa ctatgctgaa aagtacaaag tagtacagta tattgttatg tacatatcat
6360
tgtaataca gtccctggcat tctgtacata tatgtattac atttctacat ttttaatac
6420

cacatgggct tatgcattaa gtttaattgt gataaaatttg tgctgttcca gtatatgcaa
 6480
 tacacttttaa tgtttttatc ttgtacataa aaatgtgcaa tatggagatg tatacagttc
 6540
 ttactatatt aggtttataa acagtttttaa gaatttcacg cttttgcaa aatgggtggag
 6600
 tatgtaattg gtaaatcata aatcctgtgg tgaatgtggg tgtactttaa agctgtcacc
 6660
 atgttatatt ttcttttaag acattaattt agtaatttta tatttgggaa aataaagggt
 6720
 ttttaatttta ttttaactgga atcaactgcc tgctgtaatt aaacattctg taccacatct
 6780
 gtattaaaaa gacattgctg accattaaaa aaaaaa
 6816

<210> 158

<211> 1572

<212> PRT

<213> Homo sapiens

<400> 158

Ala	Ser	Gly	Asn	Leu	Asp	Lys	Asn	Ala	Arg	Phe	Ser	Ala	Ile	Tyr	Arg
1			5						10					15	
Gln	Asp	Ser	Asn	Lys	Leu	Ser	Asn	Asp	Met	Leu	Lys	Leu	Leu	Ala	
			20					25				30			
Asp	Phe	Arg	Lys	Pro	Glu	Lys	Met	Ala	Lys	Leu	Pro	Val	Ile	Leu	Gly
			35				40					45			
Asn	Leu	Asp	Ile	Thr	Ile	Asp	Asn	Val	Ser	Ser	Asp	Phe	Pro	Asn	Tyr
			50				55				60				
Val	Asn	Ser	Ser	Tyr	Ile	Pro	Thr	Lys	Gln	Phe	Glu	Thr	Cys	Ser	Lys
							70			75			80		
Thr	Pro	Ile	Thr	Phe	Glu	Val	Glu	Glu	Phe	Val	Pro	Cys	Ile	Pro	Lys
							85			90			95		
His	Thr	Gln	Pro	Tyr	Thr	Ile	Tyr	Thr	Asn	His	Leu	Tyr	Val	Tyr	Pro
			100				105					110			
Lys	Tyr	Leu	Lys	Tyr	Asp	Ser	Gln	Lys	Ser	Phe	Ala	Lys	Ala	Arg	Asn
			115				120				125				
Ile	Ala	Ile	Cys	Ile	Glu	Phe	Lys	Asp	Ser	Asp	Glu	Glu	Asp	Ser	Gln
			130				135				140				
Pro	Leu	Lys	Cys	Ile	Tyr	Gly	Arg	Pro	Gly	Gly	Pro	Val	Phe	Thr	Arg
			145				150			155				160	
Ser	Ala	Phe	Ala	Ala	Val	Leu	His	His	His	Gln	Asn	Pro	Glu	Phe	Tyr
							165			170			175		
Asp	Glu	Ile	Lys	Ile	Glu	Leu	Pro	Thr	Gln	Leu	His	Glu	Lys	His	His
							180			185			190		
Leu	Leu	Leu	Thr	Phe	Phe	His	Val	Ser	Cys	Asp	Asn	Ser	Ser	Lys	Gly
							195			200		205			
Ser	Thr	Lys	Lys	Arg	Asp	Val	Val	Glu	Thr	Gln	Val	Gly	Tyr	Ser	Trp
							210			215		220			
Leu	Pro	Leu	Leu	Lys	Asp	Gly	Arg	Val	Val	Thr	Ser	Glu	Gln	His	Ile
							225			230		235		240	
Pro	Val	Ser	Ala	Asn	Leu	Pro	Ser	Gly	Tyr	Leu	Gly	Tyr	Gln	Glu	Leu
							245			250			255		
Gly	Met	Gly	Arg	His	Tyr	Gly	Pro	Glu	Ile	Lys	Trp	Val	Asp	Gly	Gly

260 265 270
 Lys Pro Leu Leu Lys Ile Ser Thr His Leu Val Ser Thr Val Tyr Thr
 275 280 285
 Gln Asp Gln His Leu His Asn Phe Phe Gln Tyr Cys Gln Lys Thr Glu
 290 295 300
 Ser Gly Ala Gln Ala Leu Gly Asn Glu Leu Val Lys Tyr Leu Lys Ser
 305 310 315 320
 Leu His Ala Met Glu Gly His Val Met Ile Ala Phe Leu Pro Thr Ile
 325 330 335
 Leu Asn Gln Leu Phe Arg Val Leu Thr Arg Ala Thr Gln Glu Glu Val
 340 345 350
 Ala Val Asn Val Thr Arg Val Ile Ile His Val Val Ala Gln Cys His
 355 360 365
 Glu Glu Gly Leu Glu Ser His Leu Arg Ser Tyr Val Lys Tyr Ala Tyr
 370 375 380
 Lys Ala Glu Pro Tyr Val Ala Ser Glu Tyr Lys Thr Val His Glu Glu
 385 390 395 400
 Leu Thr Lys Ser Met Thr Thr Ile Leu Lys Pro Ser Ala Asp Phe Leu
 405 410 415
 Thr Ser Asn Lys Leu Leu Lys Tyr Ser Trp Phe Phe Asp Val Leu
 420 425 430
 Ile Lys Ser Met Ala Gln His Leu Ile Glu Asn Ser Lys Val Lys Leu
 435 440 445
 Leu Arg Asn Gln Arg Phe Pro Ala Ser Tyr His His Ala Val Glu Thr
 450 455 460
 Val Val Asn Met Leu Met Pro His Ile Thr Gln Lys Phe Arg Asp Asn
 465 470 475 480
 Pro Glu Ala Ser Lys Asn Ala Asn His Ser Leu Ala Val Phe Ile Lys
 485 490 495
 Arg Cys Phe Thr Phe Met Asp Arg Gly Phe Val Phe Lys Gln Ile Asn
 500 505 510
 Asn Tyr Ile Ser Cys Phe Ala Pro Gly Asp Pro Lys Thr Leu Phe Glu
 515 520 525
 Tyr Lys Phe Glu Phe Leu Arg Val Val Cys Asn His Glu His Tyr Ile
 530 535 540
 Pro Leu Asn Leu Pro Met Pro Phe Gly Lys Gly Arg Ile Gln Arg Tyr
 545 550 555 560
 Gln Asp Leu Gln Leu Asp Tyr Ser Leu Thr Asp Glu Phe Cys Arg Asn
 565 570 575
 His Phe Leu Val Gly Leu Leu Leu Arg Glu Val Gly Thr Ala Leu Gln
 580 585 590
 Glu Phe Arg Glu Val Arg Leu Ile Ala Ile Ser Val Leu Lys Asn Leu
 595 600 605
 Leu Ile Lys His Ser Phe Asp Asp Arg Tyr Ala Ser Arg Ser His Glu
 610 615 620
 Ala Arg Ile Ala Thr Leu Tyr Leu Pro Leu Phe Gly Leu Leu Ile Glu
 625 630 635 640
 Asn Val Gln Arg Ile Asn Val Arg Asp Val Ser Pro Phe Pro Val Asn
 645 650 655
 Ala Gly Met Thr Val Lys Asp Glu Ser Leu Ala Leu Pro Ala Val Asn
 660 665 670
 Pro Leu Val Thr Pro Gln Lys Gly Ser Thr Leu Asp Asn Ser Leu His
 675 680 685
 Lys Asp Leu Leu Gly Ala Ile Ser Gly Ile Ala Ser Pro Tyr Thr Thr

```

        690                695                700
Ser Thr Pro Asn Ile Asn Ser Val Arg Asn Ala Asp Ser Arg Gly Ser
705                710                715                720
Leu Ile Ser Thr Asp Ser Gly Asn Ser Leu Pro Glu Arg Asn Ser Glu
        725                730                735
Lys Ser Asn Ser Leu Asp Lys His Gln Gln Ser Ser Thr Leu Gly Asn
        740                745                750
Ser Val Val Arg Cys Asp Lys Leu Asp Gln Ser Glu Ile Lys Ser Leu
        755                760                765
Leu Met Cys Phe Leu Tyr Ile Leu Lys Ser Met Ser Asp Asp Ala Leu
        770                775                780
Phe Thr Tyr Trp Asn Lys Ala Ser Thr Ser Glu Leu Met Asp Phe Phe
785                790                795                800
Thr Ile Ser Glu Val Cys Leu His Gln Phe Gln Tyr Met Gly Lys Arg
        805                810                815
Tyr Ile Ala Arg Thr Gly Met Met His Ala Arg Leu Gln Gln Leu Gly
        820                825                830
Ser Leu Asp Asn Ser Leu Thr Phe Asn His Ser Tyr Gly His Ser Asp
        835                840                845
Ala Asp Val Leu His Gln Ser Leu Leu Glu Ala Asn Ile Ala Thr Glu
        850                855                860
Val Cys Leu Thr Ala Leu Asp Thr Leu Ser Leu Phe Thr Leu Ala Phe
865                870                875                880
Lys Asn Gln Leu Leu Ala Asp His Gly His Asn Pro Leu Met Lys Lys
        885                890                895
Val Phe Asp Val Tyr Leu Cys Phe Leu Gln Lys His Gln Ser Glu Thr
        900                905                910
Ala Leu Lys Asn Val Phe Thr Ala Leu Arg Ser Leu Ile Tyr Lys Phe
        915                920                925
Pro Ser Thr Phe Tyr Glu Gly Arg Ala Asp Met Cys Ala Ala Leu Cys
        930                935                940
Tyr Glu Ile Leu Lys Cys Cys Asn Ser Lys Leu Ser Ser Ile Arg Thr
945                950                955                960
Glu Ala Ser Gln Leu Leu Tyr Phe Leu Met Arg Asn Asn Phe Asp Tyr
        965                970                975
Thr Gly Lys Lys Ser Phe Val Arg Thr His Leu Gln Val Ile Ile Ser
        980                985                990
Val Ser Gln Leu Ile Ala Asp Val Val Gly Ile Gly Gly Thr Arg Phe
        995                1000                1005
Gln Gln Ser Leu Ser Ile Ile Asn Asn Cys Ala Asn Ser Asp Arg Leu
        1010                1015                1020
Ile Lys His Thr Ser Phe Ser Ser Asp Val Lys Asp Leu Thr Lys Arg
1025                1030                1035                1040
Ile Arg Thr Val Leu Met Ala Thr Ala Gln Met Lys Glu His Glu Asn
        1045                1050                1055
Asp Pro Glu Met Leu Val Asp Leu Gln Tyr Ser Leu Ala Lys Ser Tyr
        1060                1065                1070
Ala Ser Thr Pro Glu Leu Arg Lys Thr Trp Leu Asp Ser Met Ala Arg
        1075                1080                1085
Ile His Val Lys Asn Gly Asp Leu Ser Glu Ala Ala Met Cys Tyr Val
        1090                1095                1100
His Val Thr Ala Leu Val Ala Glu Tyr Leu Thr Arg Lys Glu Ala Val
1105                1110                1115                1120
Gln Trp Glu Pro Pro Leu Leu Pro His Ser His Ser Ala Cys Leu Arg

```


1125 1130 1135
 Arg Ser Arg Gly Val Phe Arg Gln Gly Cys Thr Ala Phe Arg Val
 1140 1145 1150
 Ile Thr Pro Asn Ile Asp Glu Glu Ala Ser Met Met Glu Asp Val Gly
 1155 1160 1165
 Met Gln Asp Val His Phe Asn Glu Asp Val Leu Met Glu Leu Leu Glu
 1170 1175 1180
 Gln Cys Ala Asp Gly Leu Trp Lys Ala Glu Arg Tyr Glu Leu Ile Ala
 1185 1190 1195 1200
 Asp Ile Tyr Lys Leu Ile Ile Pro Ile Tyr Glu Lys Arg Arg Asp Phe
 1205 1210 1215
 Glu Arg Leu Ala His Leu Tyr Asp Thr Leu His Arg Ala Tyr Ser Lys
 1220 1225 1230
 Val Thr Glu Val Met His Ser Gly Arg Arg Leu Leu Gly Thr Tyr Phe
 1235 1240 1245
 Arg Val Ala Phe Phe Gly Gln Ala Ala Gln Tyr Gln Phe Thr Asp Ser
 1250 1255 1260
 Glu Thr Asp Val Glu Gly Phe Phe Glu Asp Glu Asp Gly Lys Glu Tyr
 1265 1270 1275 1280
 Ile Tyr Lys Glu Pro Lys Leu Thr Pro Leu Ser Glu Ile Ser Gln Arg
 1285 1290 1295
 Leu Leu Lys Leu Tyr Ser Asp Lys Phe Gly Ser Glu Asn Val Lys Met
 1300 1305 1310
 Ile Gln Asp Ser Gly Lys Val Asn Pro Lys Asp Leu Asp Ser Lys Tyr
 1315 1320 1325
 Ala Tyr Ile Gln Val Thr His Val Ile Pro Phe Phe Asp Glu Lys Glu
 1330 1335 1340
 Leu Gln Glu Arg Lys Thr Glu Phe Glu Arg Ser His Asn Ile Arg Arg
 1345 1350 1355 1360
 Phe Met Phe Glu Met Pro Phe Thr Gln Thr Gly Lys Arg Gln Gly Gly
 1365 1370 1375
 Val Glu Glu Gln Cys Lys Arg Arg Thr Ile Leu Thr Ala Ile His Cys
 1380 1385 1390
 Phe Pro Tyr Val Lys Lys Arg Ile Pro Val Met Tyr Gln His His Thr
 1395 1400 1405
 Asp Leu Asn Pro Ile Glu Val Ala Ile Asp Glu Met Ser Lys Lys Val
 1410 1415 1420
 Ala Glu Leu Arg Gln Leu Cys Ser Ser Ala Glu Val Asp Met Ile Lys
 1425 1430 1435 1440
 Leu Gln Leu Lys Leu Gln Gly Ser Val Ser Val Gln Val Asn Ala Gly
 1445 1450 1455
 Pro Leu Ala Tyr Ala Arg Ala Phe Leu Asp Asp Thr Asn Thr Lys Arg
 1460 1465 1470
 Tyr Pro Asp Asn Lys Val Lys Leu Leu Lys Glu Val Phe Arg Gln Phe
 1475 1480 1485
 Val Glu Ala Cys Gly Gln Ala Leu Ala Val Asn Glu Arg Leu Ile Lys
 1490 1495 1500
 Glu Asp Gln Leu Glu Tyr Gln Glu Glu Met Lys Ala Asn Tyr Arg Glu
 1505 1510 1515 1520
 Met Ala Lys Glu Leu Ser Glu Ile Met His Glu Gln Ile Cys Pro Leu
 1525 1530 1535
 Glu Glu Lys Thr Ser Val Leu Pro Asn Ser Leu His Ile Phe Asn Ala
 1540 1545 1550
 Ile Ser Gly Thr Pro Thr Ser Thr Met Val His Gly Met Thr Ser Ser


```

<400> 161
nnacgcgtac gtctttcggc cgaagaagga acgtgggcag gggcctcctt cgctggccgc
60
cgcgcttgcc tgcagcgcac gatgaagggc gacgacagca gcaagatcac ccacaagatc
120
gccccggcga agcgcgaggg cgcggtatgg tggagctttg agtacttccc gccgcgcacg
180
ccgcagggca tgcagaattt gtatgacctg atcgagcgca tgagtcagct gggccccgcg
240
tttgtggaca ttacgtggaa tgcggggggc cggacgtcgg atatgacgac gcagctgggc
300
aagacggtgc atgcgtactt tgggtgcgag acgtgcatgc atctgacgtg c
351

```

```

<210> 162
<211> 117
<212> PRT
<213> Homo sapiens

```

```

<400> 162
Xaa Arg Val Arg Leu Ser Ala Glu Glu Gly Thr Trp Ala Gly Ala Ser
1 5 10 15
Phe Ala Gly Arg Arg Ala Trp Leu Ala Ala Thr Met Lys Gly Asp Asp
20 25 30
Ser Ser Lys Ile Thr His Lys Ile Ala Arg Ala Lys Arg Glu Gly Arg
35 40 45
Val Trp Trp Ser Phe Glu Tyr Phe Pro Pro Arg Thr Pro Gln Gly Met
50 55 60
Gln Asn Leu Tyr Asp Arg Ile Glu Arg Met Ser Gln Leu Gly Pro Glu
65 70 75 80
Phe Val Asp Ile Thr Trp Asn Ala Gly Gly Arg Thr Ser Asp Met Thr
85 90 95
Thr Gln Leu Val Lys Thr Val His Ala Tyr Phe Gly Val Glu Thr Cys
100 105 110
Met His Leu Thr Cys
115

```

```

<210> 163
<211> 360
<212> DNA
<213> Homo sapiens

```

```

<400> 163
gcgtgctcca tcggcacctt gcagatgggc gaattcgctg aaaacgtcgc cggtgccgtc
60
gacacctaca ccctgcgtca gcccatcggc gtatgcgcag gcatcactcc gtccaacttc
120
ccggcgatga ttccactgtg gatgttcccg atggcgattg cctgcggtaa cactttcgtg
180
ctcaaacctg ccgaacaaga ccctctgtcg acgatgctgc tggtagaact ggcgctggaa
240
gcgggtgtgc gcggcggcgt gctcaacgtg gtgcacggcg gcaaggatgt ggtggatgcg
300

```

ctgtgcaccc ataaagatat caaggcagtt tctttcgtcg gttcgaccgc cgttggtacc
360

<210> 164

<211> 120

<212> PRT

<213> Homo sapiens

<400> 164

Ala	Cys	Ser	Ile	Gly	Thr	Leu	Gln	Met	Gly	Glu	Phe	Ala	Glu	Asn	Val
1			5						10					15	
Ala	Gly	Gly	Val	Asp	Thr	Tyr	Thr	Leu	Arg	Gln	Pro	Ile	Gly	Val	Cys
			20					25					30		
Ala	Gly	Ile	Thr	Pro	Phe	Asn	Phe	Pro	Ala	Met	Ile	Pro	Leu	Trp	Met
		35				40						45			
Phe	Pro	Met	Ala	Ile	Ala	Cys	Gly	Asn	Thr	Phe	Val	Leu	Lys	Pro	Ser
	50				55						60				
Glu	Gln	Asp	Pro	Leu	Ser	Thr	Met	Leu	Leu	Val	Glu	Leu	Ala	Leu	Glu
65				70					75					80	
Ala	Gly	Val	Pro	Ala	Gly	Val	Leu	Asn	Val	Val	His	Gly	Gly	Lys	Asp
				85				90				95			
Val	Val	Asp	Ala	Leu	Cys	Thr	His	Lys	Asp	Ile	Lys	Ala	Val	Ser	Phe
		100						105					110		
Val	Gly	Ser	Thr	Ala	Val	Gly	Thr								
		115				120									

<210> 165

<211> 728

<212> DNA

<213> Homo sapiens

<400> 165

gctagcagcc ttcacccctcc tagaggggca ggctcggcga caaggggagg ggggtccccg
60
tcccagcgag ggaacggcgg ggctgggggt gccggtcgag cccggggcaa cagcttcacc
120
aagttttggca accgcaacgt cttcatgaag gacaacagct cttcttcagg cacagactcc
180
cgctcccgct cctcctccag gtccccgacg cgccacttcc gcagaagtga ctccactcca
240
gactccgaca gctectactc aggggaatgag tgtcaccctg tgggcccgcag gaaccggccc
300
cctaaggggc ggggagggtcg aggggcccac atggatcggg gccgaggcag ggcgcagcgt
360
gggaagaggc acgatctggc gccccaccaag cgcagtcgaa agaagatggc ggcgctggag
420
tgtgaggacc cggagcgaga gctgaagaag cagaagcggg cagcccgctt ccagcacgga
480
cactcccgcc gcctgcgcct cgagcccctg gtgctgcaga tgagcagcct ggagagcagt
540
ggggctgacc ctgactggca ggagctgcag atcgtgggca cctgcccctga catcaccaag
600
cactacctgc ccctcacctg tgccccggac ccgtccacgc tgcgcccctgt ggcattccct
660

gtggcagggtt ttgaaaaagt cgctgtgcat ggtcaagtgc cactggaaaag agaagcagga
 720
 ctacgcgt
 728

<210> 166
 <211> 242
 <212> PRT
 <213> Homo sapiens

<400> 166
 Ala Ser Ser Leu His Pro Pro Arg Gly Ala Gly Ser Ala Thr Arg Gly
 1 5 10 15
 Gly Gly Ala Pro Ser Gln Arg Gly Thr Pro Gly Ala Gly Gly Ala Gly
 20 25 30
 Arg Ala Arg Gly Asn Ser Phe Thr Lys Phe Gly Asn Arg Asn Val Phe
 35 40 45
 Met Lys Asp Asn Ser Ser Ser Ser Ser Thr Asp Ser Arg Ser Arg Ser
 50 55 60
 Ser Ser Arg Ser Pro Thr Arg His Phe Arg Arg Ser Asp Ser His Ser
 65 70 75 80
 Asp Ser Asp Ser Ser Tyr Ser Gly Asn Glu Cys His Pro Val Gly Arg
 85 90 95
 Arg Asn Pro Pro Pro Lys Gly Arg Gly Arg Gly Ala His Met Asp
 100 105 110
 Arg Gly Arg Gly Arg Ala Gln Arg Gly Lys Arg His Asp Leu Ala Pro
 115 120 125
 Thr Lys Arg Ser Arg Lys Lys Met Ala Ala Leu Glu Cys Glu Asp Pro
 130 135 140
 Glu Arg Glu Leu Lys Lys Gln Lys Arg Ala Ala Arg Phe Gln His Gly
 145 150 155 160
 His Ser Arg Arg Leu Arg Leu Glu Pro Leu Val Leu Gln Met Ser Ser
 165 170 175
 Leu Glu Ser Ser Gly Ala Asp Pro Asp Trp Gln Glu Leu Gln Ile Val
 180 185 190
 Gly Thr Cys Pro Asp Ile Thr Lys His Tyr Leu Arg Leu Thr Cys Ala
 195 200 205
 Pro Asp Pro Ser Thr Val Arg Pro Val Ala Phe Pro Val Ala Gly Phe
 210 215 220
 Glu Lys Val Ala Val His Gly Gln Val Pro Leu Glu Arg Glu Ala Gly
 225 230 235 240
 Leu Arg

<210> 167
 <211> 510
 <212> DNA
 <213> Homo sapiens

<400> 167
 nnacgcgtgg aaccagaact caggcccggt tgaggagtct ggtttggaac acacgggggc
 60
 gcaacacaga attgtcaggt cctgtgccgt gaccaccaac cctcggggcca tgccaggtgc
 120

tgggtgagggg caggtggctc ccgccaggcg cctgctggcc tgaccgcact ccgtccacag
 180
 gtctctcatg gcgtctctcg gctgggcttc gtgtccgcct acctctcaca gccactgctc
 240
 gatggctttg ccatgggggc ctccgtgacc atcctgacct cgcagctcaa acacctgctg
 300
 ggcgctcgga tcccgcgga ccagggggccc ggcatggtgg tcttcacatg gctgagcctg
 360
 ctgcgcggcg ccgggacagg caacgtgtgc gacgtggta ccagcacggg gtgcctggcg
 420
 gtgctgctag ccgcgaagga gctctcagac cgtaccgcac accgcctgag ggtgccgctg
 480
 cccacggagc tgctgggtcat cgtgggtggc
 510

<210> 168
 <211> 128
 <212> PRT
 <213> Homo sapiens

<400> 168
 Gly Ala Gly Gly Ser Arg Gln Ala Pro Ala Gly Leu Thr Ala Leu Arg
 1 5 10 15
 Pro Gln Val Leu Met Gly Val Leu Arg Leu Gly Phe Val Ser Ala Tyr
 20 25 30
 Leu Ser Gln Pro Leu Leu Asp Gly Phe Ala Met Gly Ala Ser Val Thr
 35 40 45
 Ile Leu Thr Ser Gln Leu Lys His Leu Leu Gly Val Arg Ile Pro Arg
 50 55 60
 His Gln Gly Pro Gly Met Val Val Leu Thr Trp Leu Ser Leu Leu Arg
 65 70 75 80
 Gly Ala Gly Gln Ala Asn Val Cys Asp Val Val Thr Ser Thr Val Cys
 85 90 95
 Leu Ala Val Leu Leu Ala Ala Lys Glu Leu Ser Asp Arg Tyr Arg His
 100 105 110
 Arg Leu Arg Val Pro Leu Pro Thr Glu Leu Leu Val Ile Val Val Ala
 115 120 125

<210> 169
 <211> 537
 <212> DNA
 <213> Homo sapiens

<400> 169
 gaattccacc gcattgctgtg tctggacgta tgtaggctgc ggtagtgtgc gaccgccggg
 60
 gccttaaagg agagcgggca tcggcggtgc agtacgagag gggaagggtg gcggatactt
 120
 attgtcgggtg cggcatcgtc catccacacc gttcagatggg tcaatggact ggtcaagcgg
 180
 gggtcacgagg ttacactggc atcagtcctc ccggcgggcc gtcactccat tgatccccga
 240
 gttcggatcc acctggcccc acacggcggg aaggcaaaat acgtcgctcaa tgccggctgg
 300

ctgcgatcag tggcggtctgg ggtgcaacct gacatcgtea acgtccacta tgcgaccggt
 360
 tatgggtctgc tcgctcgtct tgcccatatt gacgccccga cgctgctgtc ggtgtgggga
 420
 agtgacgttt acgattcccc cggggcaaat cccctcatgc gtcacatggt ccgatccaac
 480
 ttggtctcag ctactcggat cgcacgcaca agccactgca tggcgcggtg cacgcgt
 537

<210> 170

<211> 164

<212> PRT

<213> Homo sapiens

<400> 170

Cys Ala Thr Ala Gly Ala Leu Lys Glu Ser Gly His Arg Arg Cys Ser
 1 5 10 15
 Thr Arg Gly Glu Gly Val Arg Ile Leu Ile Val Gly Ala Ala Ser Ser
 20 25 30
 Ile His Thr Val Arg Trp Val Asn Gly Leu Val Lys Arg Gly His Glu
 35 40 45
 Val His Leu Ala Ser Val His Pro Ala Gly Arg His Ser Ile Asp Pro
 50 55 60
 Arg Val Arg Ile His Leu Ala Pro His Gly Gly Lys Ala Lys Tyr Val
 65 70 75 80
 Val Asn Ala Gly Trp Leu Arg Ser Val Ala Ala Gly Val Gln Pro Asp
 85 90 95
 Ile Val Asn Val His Tyr Ala Thr Gly Tyr Gly Leu Leu Ala Arg Leu
 100 105 110
 Ala His Ile Asp Ala Pro Thr Leu Leu Ser Val Trp Gly Ser Asp Val
 115 120 125
 Tyr Asp Ser Pro Arg Ala Asn Pro Leu Met Arg His Met Val Arg Ser
 130 135 140
 Asn Leu Val Ser Ala Thr Arg Ile Ala Ser Thr Ser His Cys Met Ala
 145 150 155 160
 Arg Val Thr Arg

<210> 171

<211> 391

<212> DNA

<213> Homo sapiens

<400> 171

ctagacaagc tcgcgcgggt gggcttcgac actcttggtc tacagacctt cctaactgcg
 60
 ggggagaagg agtccccgcg atggacgatt cacaaggcgc acaccgcccc tgaggctgct
 120
 ggcgtcatcc ataccgactt ccagaagggg ttcataaagg ccaggtgtgt gtctctcggc
 180
 gaccttgttg aatttggcgg cgaaaaaggag gccacggctg ctgggaagct gcggtttggag
 240
 ggcaaggagc acgttatgca ggacgggtgac gtagtggaaat tccgatttaa cgtgtagctc
 300

tggtttgata cttacttggc ttaaccgcat ctgagatccg tcatactctt ggcgtagcct
 360
 tattgggtatg aataacatgc cgtagccaaa g
 391

<210> 172
 <211> 98
 <212> PRT
 <213> Homo sapiens

<400> 172
 Leu Asp Lys Leu Ala Arg Val Gly Phe Asp Thr Leu Gly Leu Gln Thr
 1 5 10 15
 Phe Leu Thr Ala Gly Glu Lys Glu Ser Arg Ala Trp Thr Ile His Lys
 20 25 30
 Gly Asp Thr Ala Pro Glu Ala Ala Gly Val Ile His Thr Asp Phe Gln
 35 40 45
 Lys Gly Phe Ile Lys Ala Gln Val Val Ser Phe Gly Asp Leu Val Glu
 50 55 60
 Phe Gly Gly Glu Lys Glu Ala Gln Ala Ala Gly Lys Leu Arg Leu Glu
 65 70 75 80
 Gly Lys Glu Tyr Val Met Gln Asp Gly Asp Val Val Glu Phe Arg Phe
 85 90 95
 Asn Val

<210> 173
 <211> 309
 <212> DNA
 <213> Homo sapiens

<400> 173
 ccatggagtg tcccttgtgc gagcatttgg agagctatac caacacccat cctgcaggt
 60
 cccagagccg agccatttct caggagagca ggaaggagc aggccgaggg gtgctcccg
 120
 ccagccccgg aacccgaggt ctggggacgc agccgaccag ccttccttgt tggggcctct
 180
 gtttctctct cyacacaggg aagcagggag gggccgatca gcgacttagg cctgttggct
 240
 gtgggtgggt cccctgcgtt tctgggaagc cacggaccct gggatgtacc tgggtttcat
 300
 tcgcagtga
 309

<210> 174
 <211> 102
 <212> PRT
 <213> Homo sapiens

<400> 174
 Met Glu Cys Pro Leu Cys Glu His Phe Glu Ser Tyr Thr Asn Thr His
 .1 5 10 15
 Pro Cys Arg Ser Gln Ser Arg Ala Ile Ser Gln Glu Ser Arg Lys Gly

	20		25		30
Ala Gly Arg Gly Val Leu Pro Ala Ser Pro Gly Thr Arg Gly Leu Gly					
	35		40		45
Thr Gln Pro Thr Ser Pro Pro Cys Leu Gly Leu Cys Phe Leu Phe Asp					
	50		55		60
Thr Gly Lys Gln Gly Gly Ala Asp Gln Arg Leu Arg Pro Val Gly Cys					
	65		70		75
Gly Gly Val Pro Cys Val Ser Gly Lys Pro Arg Thr Leu Gly Cys Thr					
	85		90		95
Trp Val Ser Phe Ala Val					
	100				

<210> 175

<211> 8484

<212> DNA

<213> Homo sapiens

<400> 175

```

nnactttttt tttttttttt catttatgct atggagaaac cagcatggag atgtcatggg
60
agagcatgca caggccccc cctaggaggt ggtgatgtgt ttggggaggt gcttttttc
120
agggtccatcc cacacgttgt ccagttggat cctatggcag gtgggtgtgt gctttctctc
180
tcctgctttc cttctctctc cagataaggg tctgcaggat cttctgctta cgaagtgggt
240
gccaaaggact ggtggatggg tggctggaag cagcgacat gctccacagt ggaactgtct
300
gtctccacgg acttcatgta ttgttcagg atggcaaaaa cctcattgtt caagatctga
360
tacttcttga tcgggtcggc catcttcttc aggggcacat tcttaatgat tctatcttc
420
ccgtcctgcc tctgcacttt tagcaggtgg taacagaagt cgaacaggto aaagcgacgc
480
tgctgcacca gcaggacaat gatggagcaa ccagcccagt tcaagccat gccgaacaac
540
tgctcagctg tgaactcgtt ggttcccaca gggatgcagt acacgaactg catggcgctc
600
cacagccggt ggaactccac acactcatcg acgtgcata cgcattgtt gggcggtggg
660
cccgccaca tggggctctg caggtagctc cgaatgcggg tcaggatgac ctcaaactg
720
gacaggccac agcacagccg ctcttgggtc aggaggtcac cctcgcgagc aatggcgatt
780
tgctgagggg tccccagccg ctcgatcaga gggaccaggt ggagcggggc atacttggct
840
tccagacgtt tcattttggc atcaagtctc tccccctctt tcacatggac tcgcggcaag
900
atgtttctga aaggagccgc gtgcagcagg tcacacactt cttctaaga cagagctctg
960
ctcgatgagg aggcagaaga ggaatggcatt gccacttcc ctcaggctct ggaagcacgt
1020
ctgttttgag ctctgcgtac tcaatgatgt ccttcagctg gtggtggaag aactccagga
1080

```

tccctgggga gccatactca tgtcggggca agcggcatat cttgggcatac acctctatca
1140
gtgttttcac atactggaga atgggttcctt ggagcaagct cttcacaatc ttttagcagtt
1200
cctccatgac cacagcgaag ccctgataac ccaggagctt gcagatagtc ttgaaagtgt
1260
ggaggctcca cgaagtccg gtagctgcgc taaatgctgg agtagggcaa gtccaagacc
1320
ttggatccat gcagatactg agggctgtgca ttaggctgct tatctctttg aaattcctga
1380
gaaaatggta acactgtcgc aacaaaccgg ttggtagagc cgtttagaca gtagtggggc
1440
aggaagtcac agttgagctc ccagaagacg tgcagggtga tcctcccgta gggcgctgac
1500
acgttgtggg tggcctcccg gaacatggcg tcgaagccgt ccagcgtcag gtaccggctc
1560
agcagcttgt gggtcacgcg gttgatttcc aacaggccat ccagctcaac tatggaggtc
1620
aaatcttcac tttcaaatcg tccaatcgcc agttctaggg acttatacat ggctgctgag
1680
acgcgctggg tgatcagacg attgaggtct attgatctgc cgaggagctg cacatgcctc
1740
tgcttcagca gcgtctcgta gcggttagac ggcgggaggt ggcgtgctgc tccctgattc
1800
ttgcatctcg atcgtaaccg tttatcaaga agcaaaactc ctgccataac cttataatag
1860
gcaaatatct ggtctgccag cttgtagaca aactgatcaa aacacagggt cacctcagct
1920
tctatctcat cgtacaggaa ctgcttttta aacttggtca gagcatagta ggcgctgtca
1980
ttgtacagat ccaggagata gagcacgtac tccatcatgg aaggttcttt ggtttccagg
2040
atatggctcg ttagaatcca gggcatcgac atctcaatgg ggaactggat cctcctgccc
2100
atggctagct ccaggaagaa ctctcggaac cacagctgcg aaaggtcaca gcaactgctg
2160
agcgtttcac tgaaatttat caagtgagtg tagaagaatg actctcagtg aaagtctctc
2220
atggcgagga caatgggtcc atccaggctg ctccctcagg tcttcttgga gccgcttttg
2280
ctgcaatga gtgattcaag catggttctc accatgtaca gctgtgtgct ggatggcccc
2340
acagcacgcc ggggcacctt gatataaat ccacctttgg ggtcctcttc ccctctcaag
2400
catgggtcat tagggggctc tcgcctctcc tccagtcac agatggcttc tcgaattgac
2460
tgtagagcgc tgataggac attctctctc ttcctgaccg cctgcgcgag gggctcacgc
2520
agcgtcacct gggcgaaagt ctgcaatgcc gcgtagatgg tgttcctgat ggcctggtg
2580
aagacgctct ccactcctgc catgagcacc tgcaggcctt tgatcatggc gatcacctca
2640
acgaaggcaa atttttctc actgggtgtaa ttgtagcgtg tggctctctc atattctcc
2700

gcggtgccag gacagtcctt gttgcagaac ttgtctgtgg gatgaaccag cttccaagag
2760
tacacctcca tgacgtgggc gctccacttg gatagaagct gcagaccgag cagggctagg
2820
tcgaagagct cgcgatactc ctctgtctgac ttctggctgt ccagccctga gcccgtcacc
2880
acctcactgt tgctgtagcg agcgagctcg gagatgaagc ggatgtgtgc atcccggtac
2940
tgaacctcat gctcgcagat attgtactgg gggctgatgc tgctctgggt gcaegtcacat
3000
cgagatttat ttctctcgta gtggggcgctg gtcttgatat atcttgccag ttctatttgc
3060
atgtccccaa atagtgaac cacctggagt tgcttgaagt acttgctgat ttgggataag
3120
tttattcttt tcttggtcgc caacttatag atgttactga cactcccatc catcaggtag
3180
agaccaaate ccattgacttt gagaagcatg tgtttctcac tgggcgtcaa atacatcctg
3240
ttctctgtat aatccacaca cagattcaca atatctgcca ggagctcttc gtatgccagaa
3300
atcacttoga gctgctgctg cagagactgt gtgactctgt tatgattggc caggaacatg
3360
gacagattct gogattctct gatggactgt ggaatgcaca tttacgttaa aaactgtgct
3420
gccctcttgt aggcagagtg gtctgtcttg acgtgtcact tcatgttctt catagcttcc
3480
agcacacgca acatgttgat gaattgtccc agtgtgatca ggtaggcttc tgacacgaag
3540
tccttctctg gctcggcctg gcacagccgc ttcacctcgc tgcagaaccg ctcatggcc
3600
ttgcgtgaa aatacatgaa cttcatgagc ttggtagact ccggtccagc cactctact
3660
gtcttctcat agatctctac tcggttgggc tgctctgtgc atttcacctg gggaattggc
3720
cggggaacag tgcgccaggt gtacagcatg accgcatact catgtctctc ctccagcatc
3780
tcattctatg tggagtggac tgtagcctgc tcaatgtacc ttgcaatgcc cgtgacaaat
3840
gcattctcgt cctcaaagt ttgtgtcaaag tttagctggt acatgatgga ggaaggtgga
3900
ggctcgatgc atggtctgctg gtcggggagg ggaagctctt caagcaggtc caogttggac
3960
agggcatctt ccagggtgac gtgcgtggtc atggctgcag tttctgtatt ctgactggc
4020
ttctcgccgg gtccagggga ggggcctgag gctaccgcag ggcctccgcg cggggcgag
4080
gatgtctctg ccctctcgc tgcccgcca gtgcagaata cagaaactgc agccatgacc
4140
acgcagctga ccctggaaga tgccctgtcc aacgtggacc tgettgaaga gcttccctc
4200
ccgaccuagc agccatgcat cgagctcca ccttctcca tcatgtacca ggctaacttt
4260
gacacaaact ttgaggacag gaatgcattt gtcacgggca ttgcaaggta cattgagcag
4320

gctacagtcc actccagcat gaatgagatg ctggaggaag gacatgagta tgcgggtcatg
4380
ctgtacacct ggcgcagctg ctcccgggcc atcccacagg tgaaatgtaa cgagcagcct
4440
aacagagtgg aaatctacga gaaaaccgtg gaggttctgg agcctgaggt cacaaaaatg
4500
atgaatttca tgtacttcca gagaatgccc attgagcggt tctcggggga agtgaaggcg
4560
ctgtgccatg ccgagaggag gaaggacttc gtgtcagaag cctacctgat cacactgggc
4620
aaattcatca acatgttcgc tgtgctggac gagctgaaga acatgaagtg cagtgtgaag
4680
aacgaccact cagcgtacaa gagggccgct cagtttttac gtaaaatggc agatccacag
4740
tccatccagg aatcgagaa tctgtccatg ttcttgccca atcatacaa gatcacacag
4800
tctctgcagc agcagctcga agtgatttct ggctacgaag agctcctggc agatattgtg
4860
aatctgtgtg tggattacta cgagaacagg atgtatttga cggccagtga gaaacacatg
4920
cttctcaaa gtcattggatt tggctctgac ctgatggatg ggagtgtcag taacatctat
4980
aagttggatg ccaagaaaag aataaaactta tccaaaatgc acaagtactt caagcaactc
5040
caggtgggtc cgctatttgg ggacatgcaa atagaactgg caagatatat caagaccagc
5100
gcccaactac aggaaaaata atctcgatgg acgtgcacat cctccggcag cagccctcag
5160
tacaacatct gcgagcagat gatccagatc gcgaggagcc acatgcgctt catttcggag
5220
ctggcgcgct acagcaacag cgaagtggct acgggctcgg gccgccagga ggcccagaag
5280
acggacgcgg agtaccgcaa gctcttcgac ctggcgctgc agggcctcga gctgtgtctg
5340
cagtggagcg cgcacgtgat ggaagtgtat tcttgggaagc ttgtgcaccc caccgacaag
5400
tactccaaca aggaactgcc cgacagcgct gaagagtacg agcgtgccac gcgctacaac
5460
tacaccagcg aggagaagt tgccttagtg gaggtgatcg ccatgatcaa aggcctgcag
5520
gtgctgatgg gcaggatgga gagcgtgttc aaccacgcca tccggcacac cgtctatgcc
5580
gcactgcagg acttctccca ggtgacctt agggagccgc tgcggcaggc catcaagaag
5640
aagaagaacg tcattccagag tgccttcgag gccatcagga agaccgtgtg tgactgggag
5700
acggggcatg agcccttcaa tgaccacgac ttgcggggcg agaaggaccc caagagcggc
5760
ttcgacataa aagtaccag ccgcgcgctg ggacctccca gcactcagct ttactgggig
5820
agaaccatgc tagagtccct cattgcagac aaaagtgtt ccaagaaaac cttgagaagt
5880
agccttgagg ggcccacat attggacata gaaaaatttc atcgagagtc attcttctac
5940

actcacttga taaatttcag tgaacgctg cagcagtgtg gtgaccttc gcagctgtgg
6000
ttccgagagt tcttctctgga gctgaccatg ggcaggagga tccagtccc cattgagatg
6060
tcgatgccct ggatcctgac ggaccacatc ctggagacca aggaggcatc gatgatggag
6120
tacgtgtctc actccctgga cctgtacaat gacagcgccc actacgcgct caccagggttc
6180
aacaagcagt tctctgtacga cgaatttgag gccgagggtga atctatgttt tgaccaattt
6240
gtttacaagc tagcagacca gatatttgcc tattataagg ttatggcagg aagtttgctt
6300
cttgataaac ggttacgac agaatgcaag aatcaggagg ccacgatcca cctcccgccg
6360
tctaaccgct acgagacgct gctgaagcag aggcagtgtg agctcctcgg cagatcaata
6420
gacctcaatc gtctgatcac ccagcgcgtc tcagcagcca tgtataagtc cctgaagctg
6480
gcgattggac gatttgaaag tgaagatttg acctccatag ttgagctgga tggcctgttg
6540
gaaatcaacc gcattgacca caagctgctg agccggtacc tgacgctgga cggcttcgac
6600
gccatgttcc gggaggccaa ccacaacgtg tcagcgccct acgggaggat caccctgcac
6660
gtcttctggg agctcaacta tgacttctg cccaactact gctacaacgg ctctaccaac
6720
cggtttgttc ggacagtgtt accattttct caggaatttc aaagagataa gcagccta
6780
gcacagcttc agtatctgca tggatccaag gctttgaact tggcctactc cagcatttac
6840
ggcagctacc ggaacttcgt gggacctcca cactttcaag tcactctgcc gcttctcggc
6900
taccagggtg tcgccgtggt catggaggag ctgctgaagg tcgtcaagag cctgtgtcaa
6960
ggcacaatcc tgcaagtacg taagacgctg atggagggtga tgcccaagat ctgccgctg
7020
ccccggcacg agtacggctc tcttggtatc ctggagtctt tccaccacca gctgaaggac
7080
atcgtggagt acgcagagct gaagacgggtg tgcctccaga acctgcggga ggtggggaac
7140
gccatctctc tctgctgtct catcgacgag agcctgtctt tagaagaagt gtgtgacctg
7200
ctgcacgcgg ctcttttcca gaacatcttg ccgcgagtcc atgtgaaga gggggagaga
7260
cttgatgcca aaatgaaaag actagaatca aagtacgccc cgctgcatct tgtcccactg
7320
attgaaagac tggggacccc tcagcaaatt gccatcgcaa gagaggggga cctgctgaca
7380
aaggagcgcc tctgtctggg cctgtccatg ttgagggtca tccctgacac gatccggagc
7440
tttctggatg accccatctg gcgcgggcct ctgccacga atgggggtcat gcatgtggac
7500
agtggtgtgg agtttcacag actgtggagt gccatgcagt ttgtctactg cattcccgtg
7560

gggacacacg agttcacagt cgagcagtcg ttgggtgatg ggctacacgt ggctggctgt
 7620
 atgacatcgt taactcttgg gcagcagcgg cgttttgctg tgcctggattt ctgctacat
 7680
 ctacttaaag tccagaaaca tgatggcaaa gatgagatta ttaaaaaatg gcctttgaag
 7740
 aagatgggtg agagaattcg caagtccagc attctcaatg atgagatcat caccatcctg
 7800
 gataagtacc tgaagtcagg cgacggggag ggcacgccag tggagcatgt gcgctgcttc
 7860
 cagccgccca tccaccagtc cctcgccagc agctgagggc acgcgctgca ctccgtaact
 7920
 caacatggca tgcctttctc tccgtaaaat atttagtgag atttttaggg actatttttc
 7980
 agtatctctg tacctgttaa aggggggtgct ttctgatcta aaaacttaat ttataaaaat
 8040
 tgacttattt ttctagacta aaattgtata tgcttttggt aattagggaac tcttgagaat
 8100
 attggctgct gattgttgcc atcacgttcc tacaaaaatg tttttctatg ggaagtcttg
 8160
 gcagctgtgt cataaaatgc tgctgggttc attcattcat tccataagaa acttaatacc
 8220
 agcaaagcga ttaaatccct tgccagttac cattaactgt aactatttag cttttgttta
 8280
 gggatctttc tgatggtctt ttatgagcaa tcttagttct aagtcattgt tcccacccct
 8340
 tttttgtgtg ttccagaaaa tagtgaactt gattccctg cttccactaa atccagttgt
 8400
 gacaaaatct aacgtgacat cagatcgaaa gggttatagaa ataaaaactaa tgagatctaa
 8460
 aaaaaaaaaa aaaaaaaaaa aaaa
 8484

<210> 176

<211> 1393

<212> PRT

<213> Homo sapiens

<400> 176

Met	Ala	Arg	Glu	Gln	Leu	Arg	Gln	Val	Tyr	Ser	Met	Thr	Ala	Tyr	Ser
1			5				10					15			
Cys	Pro	Ser	Ser	Ser	Ile	Ser	Phe	Met	Leu	Glu	Trp	Thr	Val	Ala	Cys
		20					25					30			
Ser	Met	Tyr	Leu	Ala	Met	Pro	Val	Thr	Asn	Ala	Phe	Leu	Ser	Ser	Lys
		35				40					45				
Phe	Val	Ser	Lys	Leu	Ala	Trp	Tyr	Met	Met	Glu	Glu	Gly	Gly	Ser	
		50				55				60					
Met	His	Gly	Cys	Trp	Ser	Gly	Arg	Gly	Ser	Ser	Ser	Ser	Arg	Ser	Thr
65				70					75					80	
Leu	Asp	Arg	Ala	Ser	Ser	Arg	Val	Thr	Cys	Val	Val	Met	Ala	Ala	Val
				85					90					95	
Ser	Val	Phe	Cys	Thr	Gly	Ser	Ala	Ala	Gly	Pro	Gly	Glu	Gly	Pro	Glu
			100				105							110	
Ala	Thr	Ala	Gly	Pro	Arg	Ala	Gly	Ala	Gln	Asp	Ala	Leu	Pro	Arg	Ser

115	120	125
Ala Ala Pro Pro Val Gln Asn Thr Glu Thr Ala Ala Met Thr Thr His		
130	135	140
Val Thr Leu Glu Asp Ala Leu Ser Asn Val Asp Leu Leu Glu Glu Leu		
145	150	155
Pro Leu Pro Asp Gln Gln Pro Cys Ile Glu Pro Pro Ser Ser Ile		
165	170	175
Met Tyr Gln Ala Asn Phe Asp Thr Asn Phe Glu Asp Arg Asn Ala Phe		
180	185	190
Val Thr Gly Ile Ala Arg Tyr Ile Glu Gln Ala Thr Val His Ser Ser		
195	200	205
Met Asn Glu Met Leu Glu Glu Gly His Glu Tyr Ala Val Met Leu Tyr		
210	215	220
Thr Trp Arg Ser Cys Ser Arg Ala Ile Pro Gln Val Lys Cys Asn Glu		
225	230	235
Gln Pro Asn Arg Val Glu Ile Tyr Glu Lys Thr Val Glu Val Leu Glu		
245	250	255
Pro Glu Val Thr Lys Leu Met Asn Phe Met Tyr Phe Gln Arg Asn Ala		
260	265	270
Ile Glu Arg Phe Cys Gly Glu Val Arg Arg Leu Cys His Ala Glu Arg		
275	280	285
Arg Lys Asp Phe Val Ser Glu Ala Tyr Leu Ile Thr Leu Gly Lys Phe		
290	295	300
Ile Asn Met Phe Ala Val Leu Asp Glu Leu Lys Asn Met Lys Cys Ser		
305	310	315
Val Lys Asn Asp His Ser Ala Tyr Lys Arg Ala Ala Gln Phe Leu Arg		
325	330	335
Lys Met Ala Asp Pro Gln Ser Ile Gln Glu Ser Gln Asn Leu Ser Met		
340	345	350
Phe Leu Ala Asn His Asn Lys Ile Thr Gln Ser Leu Gln Gln Leu		
355	360	365
Glu Val Ile Ser Gly Tyr Glu Glu Leu Leu Ala Asp Ile Val Asn Leu		
370	375	380
Cys Val Asp Tyr Tyr Glu Asn Arg Met Tyr Leu Thr Pro Ser Glu Lys		
385	390	395
His Met Leu Leu Lys Val Met Gly Phe Gly Leu Tyr Leu Met Asp Gly		
405	410	415
Ser Val Ser Asn Ile Tyr Lys Leu Asp Ala Lys Lys Arg Ile Asn Leu		
420	425	430
Ser Lys Ile Asp Lys Tyr Phe Lys Gln Leu Gln Val Val Pro Leu Phe		
435	440	445
Gly Asp Met Gln Ile Glu Leu Ala Arg Tyr Ile Lys Thr Ser Ala His		
450	455	460
Tyr Glu Glu Asn Lys Ser Arg Trp Thr Cys Thr Ser Ser Gly Ser Ser		
465	470	475
Pro Gln Tyr Asn Ile Cys Glu Gln Met Ile Gln Ile Arg Glu Asp His		
485	490	495
Met Arg Phe Ile Ser Glu Leu Ala Arg Tyr Ser Asn Ser Glu Val Val		
500	505	510
Thr Gly Ser Gly Arg Gln Glu Ala Gln Lys Thr Asp Ala Glu Tyr Arg		
515	520	525
Lys Leu Phe Asp Leu Ala Leu Gln Gly Leu Gln Leu Leu Ser Gln Trp		
530	535	540
Ser Ala His Val Met Glu Val Tyr Ser Trp Lys Leu Val His Pro Thr		

545 550 555 560
 Asp Lys Tyr Ser Asn Lys Asp Cys Pro Asp Ser Ala Glu Glu Tyr Glu
 565 570 575
 Arg Ala Thr Arg Tyr Asn Tyr Thr Ser Glu Glu Lys Phe Ala Leu Val
 580 585 590
 Glu Val Ile Ala Met Ile Lys Gly Leu Gln Val Leu Met Gly Arg Met
 595 600 605
 Glu Ser Val Phe Asn His Ala Ile Arg His Thr Val Tyr Ala Ala Leu
 610 615 620
 Gln Asp Phe Ser Gln Val Thr Leu Arg Glu Pro Leu Arg Gln Ala Ile
 625 630 635 640
 Lys Lys Lys Lys Asn Val Ile Gln Ser Val Leu Gln Ala Ile Arg Lys
 645 650 655
 Thr Val Cys Asp Trp Glu Thr Gly His Glu Pro Phe Asn Asp Pro Ala
 660 665 670
 Leu Arg Gly Glu Lys Asp Pro Lys Ser Gly Phe Asp Ile Lys Val Pro
 675 680 685
 Arg Arg Ala Val Gly Pro Ser Ser Thr Gln Leu Tyr Met Val Arg Thr
 690 695 700
 Met Leu Glu Ser Leu Ile Ala Asp Lys Ser Gly Ser Lys Lys Thr Leu
 705 710 715 720
 Arg Ser Ser Leu Glu Gly Pro Thr Ile Leu Asp Ile Glu Lys Phe His
 725 730 735
 Arg Glu Ser Phe Phe Tyr Thr His Leu Ile Asn Phe Ser Glu Thr Leu
 740 745 750
 Gln Gln Cys Cys Asp Leu Ser Gln Leu Trp Phe Arg Glu Phe Phe Leu
 755 760 765
 Glu Leu Thr Met Gly Arg Arg Ile Gln Phe Pro Ile Glu Met Ser Met
 770 775 780
 Pro Trp Ile Leu Thr Asp His Ile Leu Glu Thr Lys Glu Ala Ser Met
 785 790 795 800
 Met Glu Tyr Val Leu Tyr Ser Leu Asp Leu Tyr Asn Asp Ser Ala His
 805 810 815
 Tyr Ala Leu Thr Arg Phe Asn Lys Gln Phe Leu Tyr Asp Glu Ile Glu
 820 825 830
 Ala Glu Val Asn Leu Cys Phe Asp Gln Phe Val Tyr Lys Leu Ala Asp
 835 840 845
 Gln Ile Phe Ala Tyr Tyr Lys Val Met Ala Gly Ser Leu Leu Leu Asp
 850 855 860
 Lys Arg Leu Arg Ser Glu Cys Lys Asn Gln Gly Ala Thr Ile His Leu
 865 870 875 880
 Pro Pro Ser Asn Arg Tyr Glu Thr Leu Leu Lys Gln Arg His Val Gln
 885 890 895
 Leu Leu Gly Arg Ser Ile Asp Leu Asn Arg Leu Ile Thr Gln Arg Val
 900 905 910
 Ser Ala Ala Met Tyr Lys Ser Leu Glu Leu Ala Ile Gly Arg Phe Glu
 915 920 925
 Ser Glu Asp Leu Thr Ser Ile Val Glu Leu Asp Gly Leu Leu Glu Ile
 930 935 940
 Asn Arg Met Thr His Lys Leu Leu Ser Arg Tyr Leu Thr Leu Asp Gly
 945 950 955 960
 Phe Asp Ala Met Phe Arg Glu Ala Asn His Asn Val Ser Ala Pro Tyr
 965 970 975
 Gly Arg Ile Thr Leu His Val Phe Trp Glu Leu Asn Tyr Asp Phe Leu

980 985 990
 Pro Asn Tyr Cys Tyr Asn Gly Ser Thr Asn Arg Phe Val Arg Thr Val
 995 1000 1005
 Leu Pro Phe Ser Gln Glu Phe Gln Arg Asp Lys Gln Pro Asn Ala Gln
 1010 1015 1020
 Pro Gln Tyr Leu His Gly Ser Lys Ala Leu Asn Leu Ala Tyr Ser Ser
 1025 1030 1035 1040
 Ile Tyr Gly Ser Tyr Arg Asn Phe Val Gly Pro Pro His Phe Gln Val
 1045 1050 1055
 Ile Cys Arg Leu Leu Gly Tyr Gln Gly Ile Ala Val Val Met Glu Glu
 1060 1065 1070
 Leu Leu Lys Val Val Lys Ser Leu Leu Gln Gly Thr Ile Leu Gln Tyr
 1075 1080 1085
 Val Lys Thr Leu Met Glu Val Met Pro Lys Ile Cys Arg Leu Pro Arg
 1090 1095 1100
 His Glu Tyr Gly Ser Pro Gly Ile Leu Glu Phe Phe His His Gln Leu
 1105 1110 1115 1120
 Lys Asp Ile Val Glu Tyr Ala Glu Leu Lys Thr Val Cys Phe Gln Asn
 1125 1130 1135
 Leu Arg Glu Val Gly Asn Ala Ile Leu Phe Cys Leu Leu Ile Glu Gln
 1140 1145 1150
 Ser Leu Ser Leu Glu Glu Val Cys Asp Leu Leu His Ala Ala Pro Phe
 1155 1160 1165
 Gln Asn Ile Leu Pro Arg Val His Val Lys Glu Gly Glu Arg Leu Asp
 1170 1175 1180
 Ala Lys Met Lys Arg Leu Glu Ser Lys Tyr Ala Pro Leu His Leu Val
 1185 1190 1195 1200
 Pro Leu Ile Glu Arg Leu Gly Thr Pro Gln Gln Ile Ala Ile Ala Arg
 1205 1210 1215
 Glu Gly Asp Leu Leu Thr Lys Glu Arg Leu Cys Cys Gly Leu Ser Met
 1220 1225 1230
 Phe Glu Val Ile Leu Thr Arg Ile Arg Ser Phe Leu Asp Asp Pro Ile
 1235 1240 1245
 Trp Arg Gly Pro Leu Pro Ser Asn Gly Val Met His Val Asp Glu Cys
 1250 1255 1260
 Val Glu Phe His Arg Leu Trp Ser Ala Met Gln Phe Val Tyr Cys Ile
 1265 1270 1275 1280
 Pro Val Gly Thr His Glu Phe Thr Val Glu Gln Cys Phe Gly Asp Gly
 1285 1290 1295
 Leu His Trp Ala Gly Cys Met Ile Ile Val Leu Leu Gly Gln Gln Arg
 1300 1305 1310
 Arg Phe Ala Val Leu Asp Phe Cys Tyr His Leu Leu Lys Val Gln Lys
 1315 1320 1325
 His Asp Gly Lys Asp Glu Ile Ile Lys Asn Val Pro Leu Lys Lys Met
 1330 1335 1340
 Val Glu Arg Ile Arg Lys Phe Gln Ile Leu Asn Asp Glu Ile Ile Thr
 1345 1350 1355 1360
 Ile Leu Asp Lys Tyr Leu Lys Ser Gly Asp Gly Glu Gly Thr Pro Val
 1365 1370 1375
 Glu His Val Arg Cys Phe Gln Pro Pro Ile His Gln Ser Leu Ala Ser
 1380 1385 1390
 Ser

<210> 177
 <211> 417
 <212> DNA
 <213> Homo sapiens

<400> 177
 acgcgtgatg tcacactgcc tctgccgctg ggtcctaatt cgattgcaac caccatggct
 60
 gcagttcgtg gcgcgcatag tttctggcat gcttcgcgca tcctggagac cgatccccgc
 120
 gctgccgtga aacgcctaa aaatgtgaag cgattgccca aagccgtgtc cgtggagcaa
 180
 atgcaaaagc tccttgccat acccagtctt aagactccta ccggcctcgc taatcgagcg
 240
 atacttgagt tcttatatgc taccggcgcg cgcgtgagcg agatgctggc aacagacctg
 300
 gacgatatac acctgggcga aaaacccgcg gatgaaaacg gggaatctat tgcacttccc
 360
 gggatatgtg gccttttttg aaaggagggt aaagagcggt tagtcccttt gggatcc
 417

<210> 178
 <211> 139
 <212> PRT
 <213> Homo sapiens

<400> 178
 Thr Arg Asp Val Thr Leu Pro Leu Pro Leu Gly Pro Asn Ser Ile Ala
 1 5 10 15
 Arg Thr Met Ala Ala Val Arg Gly Ala His Ser Phe Trp His Ala Ser
 20 25 30
 Arg Ile Leu Glu Thr Asp Pro Ala Ala Ala Val Lys Pro Pro Lys Asn
 35 40 45
 Val Lys Arg Leu Pro Lys Ala Val Ser Val Glu Gln Met Gln Lys Leu
 50 55 60
 Leu Ala Ile Pro Ser Leu Lys Thr Pro Thr Gly Leu Arg Asn Arg Ala
 65 70 75 80
 Ile Leu Glu Phe Leu Tyr Ala Thr Gly Ala Arg Val Ser Glu Met Leu
 85 90 95
 Ala Thr Asp Leu Asp Asp Ile His Leu Gly Glu Lys Pro Arg Asp Glu
 100 105 110
 Asn Gly Glu Ser Ile Ala Leu Pro Gly Tyr Val Arg Leu Phe Gly Lys
 115 120 125
 Gly Gly Lys Glu Arg Leu Val Pro Leu Gly Ser
 130 135

<210> 179
 <211> 362
 <212> DNA
 <213> Homo sapiens

<400> 179
 acgcgtcgaa ggtgccggtg ggggcgatca ataacatcgc gcaatccctg gaagagcctc
 60

aggtgattgc cgtgggttg atggtggaag atccccgcatc cccaagaato cggaattcgc
 120
 ccattggggc gggcagcccg aatccaaaat gtcggggcac gcccagtggt agtatggtaa
 180
 ggggcccggc ccatgtgttg nggcagcata cggatggaag tgctgggcga ggcctctggg
 240
 ttgccggcag agcaactggg gcagctcaag cggggcgggg tgatcgagca gttggattga
 300
 gcaatggcgg ccgcgaagcc cgccatttac cttgatgact gtttagcgcg cggattcttt
 360
 aa
 362

<210> 180
 <211> 108
 <212> PRT
 <213> Homo sapiens

<400> 180
 Met Ala Gly Phe Ala Ala Ala Ile Ala Gln Ser Asn Cys Ser Ile Thr
 1 5 10 15
 Pro Pro Ala Leu Ser Cys Pro Ser Cys Ser Ala Gly Lys Pro Arg Arg
 20 25 30
 Ser Pro Ser Thr Ser Ile Arg Met Leu Pro Pro Thr Ser Val Pro Ala
 35 40 45
 Pro Tyr His Thr Pro Thr Gly Arg Ala Pro Thr Phe Trp Ile Arg Ala
 50 55 60
 Ala Arg Pro Asn Gly Glu Phe Pro Asp Ser Trp Gly Cys Gly Ile Phe
 65 70 75 80
 His His Gln Pro Thr Gly Asn His Leu Arg Leu Phe Gln Gly Leu Arg
 85 90 95
 Asp Val Ile Asp Arg Pro His Arg His Leu Arg Arg
 100 105

<210> 181
 <211> 297
 <212> DNA
 <213> Homo sapiens

<400> 181
 gcgttgatca tgccgaccc aggtttgatc atgctgggtac gccgtcactt cccgtgcatg
 60
 ccgattccat tgctggatca ggccaatacg gtgaattggg ccagcgctga gttctggcaa
 120
 cagcaaggta tctgccgggt aatcctgtcg cggaattgt cactggaaga aatcggcgaa
 180
 atccgccaac aggtgcccgc catggagctg gaagtgtttg tgcacgggtc cctgtacatg
 240
 gcctattccg ggcgctgttt gttgtccggc tatatgaaca agcgcgatgc caaccaa
 297

<210> 182
 <211> 99
 <212> PRT

<213> Homo sapiens

<400> 182

Ala Leu Ile Met Ser Asp Pro Gly Leu Ile Met Leu Val Arg Arg His
 1 5 10 15
 Phe Pro Cys Met Pro Ile His Leu Ser Val Gln Ala Asn Thr Val Asn
 20 25 30
 Trp Ala Ser Val Glu Phe Trp Gln Gln Gly Ile Cys Arg Val Ile
 35 40 45
 Leu Ser Arg Glu Leu Ser Leu Glu Glu Ile Gly Glu Ile Arg Gln Gln
 50 55 60
 Val Pro Ala Met Glu Leu Glu Val Phe Val His Gly Ala Leu Tyr Met
 65 70 75 80
 Ala Tyr Ser Gly Arg Cys Leu Leu Ser Gly Tyr Met Asn Lys Arg Asp
 85 90 95
 Ala Asn Gln

<210> 183

<211> 351

<212> DNA

<213> Homo sapiens

<400> 183

cgggacgtca ccatgaagcc gaccggctcg ggggatgtgg cgaacaaggt catcacccat
 60
 attccgttta acatcgtctc ccaggcgact catccattcc ttcgtacctt ggcagatgtc
 120
 aagcgcattc ctttggcgac cgacgggctc ggccaccagg tcttgctcaa gggctaccag
 180
 gccgagggcc acgactacgc acaccccgac tacggcgcca acgtctccca cegtgcggcg
 240
 gggatgaagg attcgcagaa gtcaccagag tcgggcaggc agtggaaacac cgatttcggc
 300
 attcacgtca acctggtgga gtctatctc gaggcgaatc atttcggcga c
 351

<210> 184

<211> 117

<212> PRT

<213> Homo sapiens

<400> 184

Arg Asp Val Thr Met Lys Pro Thr Gly Ser Gly Asp Val Ala Asn Lys
 1 5 10 15
 Val Ile Thr His Ile Pro Phe Asn Ile Val Ser Gln Ala Thr His Pro
 20 25 30
 Phe Leu Arg Thr Leu Asp Asp Val Lys Arg Ile Ser Leu Ala Thr Asp
 35 40 45
 Gly Leu Gly His Gln Val Leu Leu Lys Gly Tyr Gln Ala Glu Gly His
 50 55 60
 Asp Tyr Ala His Pro Asp Tyr Gly Gly Asn Val Ser His Arg Ala Gly
 65 70 75 80
 Gly Met Lys Asp Leu Glu Lys Leu Thr Glu Ser Gly Arg Gln Trp Asn

```

      85              90              95
Thr Asp Phe Gly Ile His Val Asn Leu Val Glu Ser Tyr Pro Glu Ala
      100              105              110
Asn His Phe Gly Asp
      115

<210> 185
<211> 396
<212> DNA
<213> Homo sapiens

<400> 185
cgcggtgggtc tcagtaaaga aaatttggtg cttagaggat gcaccattag aaacacagag
60
gctgttggtg gcattgtggt ttatgcaggc catgaaacca aagcaatgct gaacaacagt
120
gggccacggtg ataagcgagc caaattagaa agaagagcaa acacagatgt cctctgggtg
180
gtcatgtctt tggtcataat gtgcttaact ggcgcagtag gtcattggaat ctggctgagc
240
aggtagaaa agatgcattt tttcaatggt cccgagcctg atggacatat catatcacca
300
ctgttggcag gattttatat gttttggacc gtgatcattt tgttacaggt cttgattcct
360
atttctctct atgtttccat cgaaattgtg aagcctt
396

<210> 186
<211> 132
<212> PRT
<213> Homo sapiens

<400> 186
Arg Val Gly Leu Ser Lys Glu Asn Leu Leu Leu Arg Gly Cys Thr Ile
  1              5              10              15
Arg Asn Thr Glu Ala Val Val Gly Ile Val Val Tyr Ala Gly His Glu
      20              25              30
Thr Lys Ala Met Leu Asn Asn Ser Gly Pro Arg Tyr Lys Arg Ser Lys
      35              40              45
Leu Glu Arg Arg Ala Asn Thr Asp Val Leu Trp Cys Val Met Leu Leu
      50              55              60
Val Ile Met Cys Leu Thr Gly Ala Val Gly His Gly Ile Trp Leu Ser
      65              70              75              80
Arg Tyr Glu Lys Met His Phe Phe Asn Val Pro Glu Pro Asp Gly His
      85              90              95
Ile Ile Ser Pro Leu Leu Ala Gly Phe Tyr Met Phe Trp Thr Val Ile
      100              105              110
Ile Leu Leu Gln Val Leu Ile Pro Ile Ser Leu Tyr Val Ser Ile Glu
      115              120              125
Ile Val Lys Leu
      130

<210> 187
<211> 423

```

<212> DNA

<213> Homo sapiens

<400> 187

cgagtgctca ccgcgtcagc cgtcatgcgt cccactgagg ctgttgctct tcggtcggca
 60
 gaacctcgac gagttcagcg gatcctggac cagcgcgagt gggctggcgt ctctgttgct
 120
 gatgagcatt gtcgtttgct tggcacggtc ggcgatcaag aggtcatcga ggctgctcgc
 180
 cgcggagatc gcagttattgc tgacgcgggtg gaaactaacg gcatactcac ggccgcggacc
 240
 gacactccgt tgctcagact ctctgcctcg accagcaacg ccaggggtgcc gttggccggtt
 300
 gtcgacgagg acttccacct catgggtgct atctctcggg tgacctctgt cgacgcgatg
 360
 tcacgagctc gcgacgaggc aggagagggg tctgtcatgt ccttgagaaa caccggaaa
 420
 ctt
 423

<210> 188

<211> 141

<212> PRT

<213> Homo sapiens

<400> 188

Arg Val Leu Thr Ala Ser Ala Val Met Arg Pro Thr Glu Ala Val Val
 1 5 10 15
 Ser Arg Ser Ala Glu Pro Arg Arg Val Gln Arg Ile Leu Asp Gln Arg
 20 25 30
 Glu Trp Ala Gly Val Phe Val Val Asp Glu His Arg Arg Leu Leu Gly
 35 40 45
 Thr Val Gly Asp Gln Glu Val Ile Glu Ala Arg Arg Gly Asp Arg
 50 55 60
 Ser Ile Ala Asp Ala Val Glu Thr Asn Gly Ile Leu Thr Ala Arg Thr
 65 70 75 80
 Asp Thr Pro Leu Ser Glu Leu Phe Ala Pro Thr Ser Asn Ala Arg Val
 85 90 95
 Pro Leu Ala Val Val Asp Glu Asp Phe His Leu Met Gly Val Ile Ser
 100 105 110
 Arg Val Thr Leu Leu Asp Ala Met Ser Arg Ala Arg Asp Glu Ala Gly
 115 120 125
 Glu Gly Ser Val Met Ser Leu Glu Asn Thr Gly Lys Leu
 130 135 140

<210> 189

<211> 429

<212> DNA

<213> Homo sapiens

<400> 189

ngatgggttta ccaacatatg cactggttcga gcggcaatat ctctctgggg gctggcagtg
 60

aaatgtttga agatgccggc gttccgggcc tcaacttgtt tcgatgccgt ggttccaccg
 120
 atttcgcga tgcggctcat cgcacgggta agaagtttcg tccagataac ccaggacaga
 180
 gcaaggata tcaggctcag aaccaggaag agcagggttc tacccagtg ccccatatag
 240
 accgcgcatg ctacggcaaa aggcgcgcgc agtgggggtcc aggacagcac ttctatggct
 300
 gaaggagcgc catcccnagc ttgccttagc cccagagcta acccagcgac cagtggacca
 360
 gcgcccacat tcagtaggaa ccttacgata atcagccctt gttttaccgc tggaatggag
 420
 ctgatttcn
 429

<210> 190
 <211> 123
 <212> PRT
 <213> Homo sapiens

<400> 190
 Met Met Gly Ala Gly Pro Leu Val Ala Gly Leu Ala Leu Gly Leu Gly
 1 5 10 15
 Glu Ala Xaa Asp Ala Leu Pro Ser Ala Met Lys Val Leu Ser Trp Thr
 20 25 30
 Pro Leu Gly Ala Pro Phe Ala Val Ala Ser Ala Val Tyr Met Gly His
 35 40 45
 Trp Gly Lys Ala Leu Leu Phe Leu Val Leu Ser Leu Ile Tyr Leu Ala
 50 55 60
 Leu Ser Trp Val Ile Trp Thr Lys Leu Leu Asn Arg Ala Met Ser Arg
 65 70 75 80
 Ile Gly Glu Ile Gly Gly Thr Thr Ala Ser Lys Gln Val Glu Ala Gly
 85 90 95
 Asn Ala Gly Ile Phe Lys His Phe Thr Ala Ser Pro Arg Gly Ala Ile
 100 105 110
 Ala Ala Arg Thr Val His Met Leu Val Asn His
 115 120

<210> 191
 <211> 4845
 <212> DNA
 <213> Homo sapiens

<400> 191
 ccgccggggg ccattggcgac actcagcttc gtcttcctgc tgctgggggc agtgtcctgg
 60
 cctccggctt ctgcctccgg ccaggagttc tggcccgac aatcgccgc cgatatcttg
 120
 tcggggggcg cttcccgag acggtatctt ctgtatgacg tcaaccccc ggaaggcttc
 180
 aacctcgca gggatgtcta tatccgaatc gcctctctcc tgaagactct gctgaagacg
 240
 gaggagtggg tgcttgctct gcctccatgg ggccgcctct atcactggca gactcctgac
 300

atccaccagg tccgattcc ctggtctgag ttttttgatc ttccaagtct caataaaaaac
 360
 atccccgtca tccagtatga gcagttcatc gcagaatctg gtgggcccctt tattgaccag
 420
 gtttacgtcc tgcaaaagta cgcagagggg tggaaagaag ggacctggga agagaagggtg
 480
 gacgacgggc cgtgtattga tcagctcctg tactcccagg acaagcacga gtactacaga
 540
 ggatggtttt gggggttatga ggagaccagg ggtctaaacg tctcctgtct gtccgtccag
 600
 ggctcagctc ccatcgtggc gccctgctg ctgagaaaca catcagcccg gtccgtgatg
 660
 ttagacagag ccgagaacct acttcacgac cactatggag ggaaagaata ctgggatacc
 720
 cgctgcagca tgggtgtttgc caggcacctg cgggagggtg gagacgagtt caggagcaga
 780
 catctcaact ccaoggacga cgcagacagg atcccccttc aggaggactg gatgaagatg
 840
 aagggtcaag tgggctccgc gctagggggc cctacctcg gagtccacct gagaagaaaa
 900
 gatttcatct ggggtcacag acaggatgta ccagctctg aaggggccgt gaggaagatc
 960
 cgcagcctca tgaagaccca cgggctggac aagggtgttg tggccacaga tgccgtcaga
 1020
 aaggaatatg aagagctaaa aaagctgtta ccgagatgg tagggtttga acccagtggtg
 1080
 gaggagctgg agctctacaa ggacggaggc gttgcgatta ttgaccagtg gatctgcgca
 1140
 cacgccaggt gcctgcccac gtcactgtcg ccgagagcgg ggtcgggtgt ctttcaaagg
 1200
 ttcttctgtc ccaagtactc ggtgtcagag cagatgggtc cctgtgttca cagtggctcat
 1260
 ttccatactg tttgcctcct cgtctgagtc tctgtagca tctgggtcag tgtttccctg
 1320
 ggctgaagtt aattgttcat ctgccccct tagttctcat gcacagaatt cctccatagc
 1380
 aggcgtgttg catagctggc ctgctctcag aacctctctt tgtgtcgcat ttcccatca
 1440
 ttcccggttt ctgccccctg ctgccccctg cctgagagt tgcccggtcc ctggactgtg
 1500
 gcatgtcctt gttgctgtgt tgttgagcat ccgtgagcgt ccccgaggcc ggagcgtgg
 1560
 gccctcgtgt gatcattctc gtggggctgc catgagcgtc cccaaggctg ggagcatggg
 1620
 ccctcgtgtg atcgttcttg tggggctgcc gtgagcgtcc ccgaggccgg gagcgtgggc
 1680
 cctcgtgtga tcattctcgt ggggctgccg tgagcgtccc agaggccggg agcgtggggc
 1740
 ctgctgtcag tcattctctt ggggttctgt tgggaggtac gcctgggctc ctgttctctc
 1800
 aaagacctgc ctgcccctct gcataggaga tgaaggctgg ggttaggggt aaacggtttg
 1860
 agttaaatgg aaaatgaaag tagagggaaat gatcttccc gtggttagca ctgtgcacac
 1920

gcgtgcgtct ctgtgggtta gtctgtctct ctctgcccg aggaatgctg agcgccctga
1980
gccggtgcct cttcacacat ctgctatttc ctgtggtggt ctgggcatgg tgataagac
2040
ccacagaggc tcgggtgat gctgtctgct ggtgtgggt cctttccct gttaagcaga
2100
caggtatgac cgctgacttc ttaggtcagg gcggaggtgt gcaggagccc agtcacgagc
2160
tcaccctctg ttctcaggtg tggccttggg attttgactg cgaccttggc ggtgctgtct
2220
ccgcagccca ggaagcctgc tgtggggagg ctctgcactg agctctcagc ctctgccct
2280
cagctgcgcg aagcgctcgg ccagctcac tgaagctgcc ctgcctccgg gccggcgcg
2340
cctgctctgg caggccctg tgtgtgggt ggtgagggtc tccccaccg tctgacccc
2400
cgacgacga tacaggcctg tgtggcctgc tggccctgtg gctctgtgtg cagcgctgtg
2460
catgttacct ttgctctgga aacatctctg ggttttgcct gttcacgaag ttcataagat
2520
gcctgtggag agccagagac cagctgcgca ggagccggag gaacgggacg gccgtgacc
2580
tgagggtctg agaaacccct ggagaagggt gtccccacca gccatacagc cgtgtgtgtg
2640
gagggggcct tgacctcogt gatgtctact gtgcctcagg ataaggaccc gccatgcct
2700
ggctagacag tgtgtgggtta tagggaatct ctcttcttc accatgtgac cccaggagg
2760
tattcgacct gcactggcgt gccctggcctg ggatttgggt acggagagga gggctccag
2820
ggggacatgc ggtgggacag gacggtggcg gctgctgact gtggggtgtg gatggggctg
2880
cagcaccagg cagagccctc caggggcgcc atgtcgaggg cacttgagcg aggggtgcc
2940
gcaagggggc ggccggctgg tgggtgctcg ggggacggcc gtgttgggtc catgtaaactg
3000
tggaactcat ttcactcagc tgctcctctc agttctccct gactctggaa cctctgtgac
3060
cccattctc cctgactctg gaacctctgt gatcctttgc aggttttttt attggcacct
3120
cagttctaac attttctttt cggattcatg aggaagaga aatcctgggg ttggacccca
3180
agacgacgta caacaggttc tcgggagacc aagagaaggc gtgtgagcaa cccaccact
3240
ggaagatcac ctactgagga ggatccctcca gggccgcctc ccggacccca caggcgcggg
3300
tggatcgagc ttctgtgcc gtggagtac cgtctactgc cagcgggagc ctggcgggac
3360
aggaccgtcc ctgcagggtt ccagggccca gaagaggccc cagcctcta gagctgggtc
3420
cgctcctcgg cgttgccagc cggcatggct gatgaagggt ctccgtgctc ctggggggcg
3480
cggttggttt tcaggcagcg tctgtgaacc cacagctcgg ttgccagcag tgcggcggtg
3540

gtgaccaga agcaggagtg tttgtcaggc tcccgtcttg gcctttccag ccacctttca
 3600
 tgtcttcata ttttaagtgc attgaggata gatgcaggcg ggtgagctgc cctccgtcag
 3660
 ggggaccog gctgacattt ccctgggagc tgggtgcaagg agaagcgta ttttaaatgt
 3720
 ctgcagagcg accagggggcc tcatgaatct ctccgttgcc ctccgcgcag caggaggctg
 3780
 cctgtgtgtt tctctctggg acgogtgcaa ggcagacctg gtgctgcaaa ggaaggggcc
 3840
 tgaggcctca gggagccccg tggagggatg acagttcagg cctactgtct ggcacgtcag
 3900
 agcactggga agtttttcag tgacgtctct ggggcactca gtggattgtc tgtaggaaac
 3960
 ttgcagctct gctcctcaca ccaggcccg ctggccaccc accctcgccc ccaactggcca
 4020
 cccctccctc ccccgagctg ccccgcccca ccctcacccc gactgccccg ccctcgcccg
 4080
 gctggccgtc cctgcctctg ccccggtctg cagggtgcaa tggggcctcc aggtctgcga
 4140
 ttcgctattg agaactagaa atgagggaagg acagttacgc taactccaaa aggctgtcta
 4200
 ggatgagctg ctttatcagg gagctccttg taccatttt acagaaatca tttttaggtc
 4260
 tttgtgccac caccacgagg ggcattctga aagagggcaa cgctagacac aagaatccgtg
 4320
 gaagggtcag cagtgcctca ggggtcctca gggtcaggga gccccctca cctctttggc
 4380
 ccgttaccct ttgtgacttt ccaccatggt gtcgtgtgac cctcagtcag gttggtgggg
 4440
 gctgagtcct cactgagcag ccactttcca catctgctag aggaacagtg acatggacac
 4500
 ctgtgacaga gagaggacag ttagtgagga gggacagaca gctcttctt tcggagcctg
 4560
 gctagtctag gacatcacct tgctgtgtct tctcaagctt ttaaaattga cctgaacgt
 4620 cctattggtgt tactcaaagc tgtgcagggt aaatgatgac atatttattc 4680
 tttttccatt tgttctagaa acagtgccct tttcatcagt tgcattttcc aggctgagag
 4740
 ctgtataaaa cattttggac tgtgaccatg taccttctt ttaagaaaa ataaactgct
 4800
 ttatggaagt tggtaaaaa aaaaaaaaaa aaaaaaaa aaaaa
 4845

<210> 192

<211> 428

<212> PRT

<213> Homo sapiens

<400> 192

Pro Pro Gly Ala Met Ala Thr Leu Ser Phe Val Phe Leu Leu Leu Gly
 1 5 10 15
 Ala Val Ser Trp Pro Pro Ala Ser Ala Ser Gly Gln Glu Phe Trp Pro
 20 25 30
 Gly Gln Ser Ala Ala Asp Ile Leu Ser Gly Ala Ala Ser Arg Arg Arg

```

      35              40              45
Tyr Leu Leu Tyr Asp Val Asn Pro Pro Glu Gly Phe Asn Leu Arg Arg
 50              55              60
Asp Val Tyr Ile Arg Ile Ala Ser Leu Leu Lys Thr Leu Leu Lys Thr
65              70              80
Glu Glu Trp Val Leu Val Leu Pro Pro Trp Gly Arg Leu Tyr His Trp
      85              90              95
Gln Ser Pro Asp Ile His Gln Val Arg Ile Pro Trp Ser Glu Phe Phe
      100              105              110
Asp Leu Pro Ser Leu Asn Lys Asn Ile Pro Val Ile Glu Tyr Glu Gln
      115              120              125
Phe Ile Ala Glu Ser Gly Gly Pro Phe Ile Asp Gln Val Tyr Val Leu
      130              135              140
Gln Ser Tyr Ala Glu Gly Trp Lys Glu Gly Thr Trp Glu Glu Lys Val
      145              150              155
Asp Glu Arg Pro Cys Ile Asp Gln Leu Leu Tyr Ser Gln Asp Lys His
      165              170              175
Glu Tyr Tyr Arg Gly Trp Phe Trp Gly Tyr Glu Glu Thr Arg Gly Leu
      180              185              190
Asn Val Ser Cys Leu Ser Val Gln Gly Ser Ala Ser Ile Val Ala Pro
      195              200              205
Leu Leu Leu Arg Asn Thr Ser Ala Arg Ser Val Met Leu Asp Arg Ala
      210              215              220
Glu Asn Leu Leu His Asp His Tyr Gly Gly Lys Glu Tyr Trp Asp Thr
      225              230              235
Arg Arg Ser Met Val Phe Ala Arg His Leu Arg Glu Val Gly Asp Glu
      245              250              255
Phe Arg Ser Arg His Leu Asn Ser Thr Asp Asp Ala Asp Arg Ile Pro
      260              265              270
Phe Gln Glu Asp Trp Met Lys Met Lys Val Lys Leu Gly Ser Ala Leu
      275              280              285
Gly Gly Pro Tyr Leu Gly Val His Leu Arg Arg Lys Asp Phe Ile Trp
      290              295              300
Gly His Arg Gln Asp Val Pro Ser Leu Glu Gly Ala Val Arg Lys Ile
      305              310              315
Arg Ser Leu Met Lys Thr His Arg Leu Asp Lys Val Phe Val Ala Thr
      325              330              335
Asp Ala Val Arg Lys Glu Tyr Glu Glu Leu Lys Lys Leu Leu Pro Glu
      340              345              350
Met Val Arg Phe Glu Pro Thr Trp Glu Glu Leu Glu Leu Tyr Lys Asp
      355              360              365
Gly Gly Val Ala Ile Ile Asp Gln Trp Ile Cys Ala His Ala Arg Cys
      370              375              380
Leu Pro Thr Ser Leu Ser Ala Glu Ser Gly Ser Gly Gly Phe Gln Arg
      385              390              395
Phe Phe Cys Pro Lys Tyr Ser Val Ser Glu Gln Met Val Ala Cys Val
      405              410              415
His Ser Gly His Phe His Thr Val Cys Leu Leu Val
      420              425

```

<210> 193

<211> 350

<212> DNA

<213> Homo sapiens

<400> 193
 gccggcgagc tggactgcgc catcatggcc gagcccttcc ccgacacccg cctgggccacg
 60
 gcgcagctgt acgacgagcc ctctgcgtgc gcgctgcggg cgtcgacccc gctggccgac
 120
 cgtgccagca tcagcccccga ggaggtcaag ggcgagacca tgttgatgtt gggcacgggc
 180
 ccttggtttc cccgggcccgc cgggtgggggt ttggcccgga tttggcgctt ttctccagcg
 240
 ccgttaaggc catacgccgc agtttcgagg gctcgtcgtt ggagaccatc aagcacatcg
 300
 tggcttcggg catggcgtag cgggtggtgcc gcagctgtcc gtgccgcgcg
 350

<210> 194
 <211> 116
 <212> PRT
 <213> Homo sapiens

<400> 194
 Ala Gly Glu Leu Asp Cys Ala Ile Met Ala Glu Pro Phe Pro Asp Thr
 1 5 10 15
 Gly Leu Ala Thr Ala Gln Leu Tyr Asp Glu Pro Phe Val Val Ala Leu
 20 25 30
 Arg Ala Ser His Pro Leu Ala Asp Arg Ala Ser Ile Ser Pro Glu Glu
 35 40 45
 Val Lys Gly Glu Thr Met Leu Met Leu Gly Thr Gly Pro Trp Phe Pro
 50 55 60
 Arg Ala Arg Gly Gly Leu Ala Arg Ile Trp Arg Val Ser Pro Ala
 65 70 75 80
 Pro Leu Arg Ala Tyr Ala Ala Val Ser Arg Ala Arg Arg Trp Arg Pro
 85 90 95
 Ser Ser Thr Ser Trp Leu Arg Ala Trp Arg Asp Gly Gly Ala Ala Ala
 100 105 110
 Val Arg Ala Ala
 115

<210> 195
 <211> 495
 <212> DNA
 <213> Homo sapiens

<400> 195
 acgcgtgaac gcgacggctt ggcgatcgga ggcgtcgccc ccgtcgttga gtgggcccgtt
 60
 gaaatgtgtc gcttcgacga aagcgagact ctgcaccgcc ttgcacatggg cgtctctgaa
 120
 ccagaacttg gcgacgattt ggccgcccgc ctgctcgatt ctcatcggtt tgctgtcatc
 180
 agcgagggat cgaactggct tgcctcgcta cccgtgatcg taggtcgcaa caccgaaacg
 240
 ttctcgagca taccagacct tgcccgcgac cggatcgaca aactgcacca gttgagccat
 300

cgcgaaatag cagcaaatcg cgagctctcg cgtgcccgcg ctgcgtcggg gcaggtgcgg
 360
 cactgccacg ggcagcgaca cctcgccaac atcgctcatga ttgacggcaa gccggtctcg
 420
 ttgcagcgga tcgaatttga tctcgatatc ggcacaacgg atgtgctgta cgatttcggg
 480
 ttccctctga tggat
 495

<210> 196

<211> 165

<212> PRT

<213> Homo sapiens

<400> 196

Thr	Arg	Glu	Arg	Asp	Gly	Leu	Ala	Ile	Gly	Gly	Val	Gly	Pro	Val	Val
1				5					10				15		
Glu	Trp	Ala	Val	Glu	Met	Val	Arg	Phe	Asp	Glu	Ser	Glu	Thr	Leu	Asp
		20						25				30			
Arg	Leu	Ala	Ser	Gly	Val	Leu	Glu	Pro	Glu	Leu	Gly	Asp	Asp	Leu	Ala
	35					40					45				
Ala	Val	Leu	Leu	Asp	Ser	His	Arg	Val	Ala	Val	Ile	Ser	Glu	Gly	Ser
	50				55					60					
Asn	Trp	Leu	Ala	Ser	Leu	Pro	Val	Ile	Val	Gly	Arg	Asn	Thr	Glu	Gln
65				70					75				80		
Phe	Arg	Ser	Ile	Pro	Asp	Leu	Ala	Arg	Asp	Arg	Ile	Asp	Lys	Leu	His
			85					90				95			
Gln	Leu	Ser	His	Arg	Glu	Ile	Ala	Arg	Asn	Arg	Glu	Leu	Leu	Arg	Ala
			100				105					110			
Arg	Ala	Ala	Ser	Gly	Gln	Val	Arg	His	Cys	His	Gly	Asp	Ala	His	Leu
	115				120						125				
Gly	Asn	Ile	Val	Met	Ile	Asp	Gly	Lys	Pro	Val	Leu	Phe	Asp	Ala	Ile
	130				135					140					
Glu	Phe	Asp	Pro	Asp	Ile	Ala	Thr	Thr	Asp	Val	Leu	Tyr	Asp	Phe	Ala
145				150					155					160	
Phe	Pro	Leu	Met	Asp											
					165										

<210> 197

<211> 402

<212> DNA

<213> Homo sapiens

<400> 197

caagcaatgc ttgacgcagt tgttgaatac ttaccagcac cgactgatatt tccagcaatc
 60
 aaaggatatca atccagatga aactgaaggt gaacgtcacg caagcgatga tgagccattc
 120
 tcttcatttag cattcaaaat tgcaactgac ccattcgttag gtaacttaac cttcttcctg
 180
 gtgtactcag gtgtaattaa ctctgggtgat acagtattaa actctgttacg tcaaaaaacg
 240
 gaacgttttg tgcgtatcgt acagatgcac gctaataaac gtgaagaaat taaagaagtt
 300

cgctgcggcg atatacgtgc agcaatcggc ttaaaagatg taactacggg tgaaccatta
 360
 tgtgctgtcg atgcaccaat cattcttgag cgtatggaat tc
 402

<210> 198
 <211> 134
 <212> PRT
 <213> Homo sapiens

<400> 198
 Gln Ala Met Leu Asp Ala Val Val Glu Tyr Leu Pro Ala Pro Thr Asp
 1 5 10 15
 Ile Pro Ala Ile Lys Gly Ile Asn Pro Asp Glu Thr Glu Gly Glu Arg
 20 25 30
 His Ala Ser Asp Asp Glu Pro Phe Ser Ser Leu Ala Phe Lys Ile Ala
 35 40 45
 Thr Asp Pro Phe Val Gly Asn Leu Thr Phe Phe Arg Val Tyr Ser Gly
 50 55 60
 Val Ile Asn Ser Gly Asp Thr Val Leu Asn Ser Val Arg Gln Lys Arg
 65 70 75 80
 Glu Arg Phe Gly Arg Ile Val Gln Met His Ala Asn Lys Arg Glu Glu
 85 90 95
 Ile Lys Glu Val Arg Ala Gly Asp Ile Ala Ala Ile Gly Leu Lys
 100 105 110
 Asp Val Thr Thr Gly Glu Pro Leu Cys Ala Val Asp Ala Pro Ile Ile
 115 120 125
 Leu Glu Arg Met Glu Phe
 130

<210> 199
 <211> 507
 <212> DNA
 <213> Homo sapiens

<400> 199
 acgcgtgaag tcgtgcatag atcgggtgtga catagagaag cctccgaccc aagctgcgta
 60
 tatcgacaaa agaccaagcg accctggacg ttctagacag aactctgcta cgaggcctga
 120
 caatagttaa atccccgaga acccagctat ggaagggttt ccagatgctc gaaggcctgt
 180
 cataccagag gttaggttaa actgtatgga gactttcgag gtgaaagtgt actcgcgggt
 240
 aaagcctgct cctaagaggg atttagatct gatagatcta tectcagatt caacctcggt
 300
 gctgaaaaa cactctatc tctcaacctc cgacagcgac tctctgttat ttgagcctct
 360
 tccctctctc agaatagtcg agagtgcga agaagaggag acgatgaacc aaggcgatga
 420
 cggccccctc ggtaaaaatg ctgcctcttc tccctccatc cccagccatc cctccgtctc
 480
 cagcctgagc acagctccgc ttgtaca
 507

<210> 200
 <211> 153
 <212> PRT
 <213> Homo sapiens

<400> 200
 Met Glu Gly Glu Glu Ala Ala Phe Leu Pro Glu Gly Pro Ser Ser Pro
 1 5 10 15
 Trp Phe Ile Val Ser Ser Ser Ser Ser Leu Ser Thr Ile Leu Arg Glu
 20 25 30
 Gly Arg Gly Ser Asn Thr Arg Glu Ser Leu Ser Glu Val Glu Ser Ile
 35 40 45
 Glu Cys Phe Ser Gly Pro Glu Val Glu Ser Glu Asp Arg Ser Ile Arg
 50 55 60
 Ser Lys Ser Ser Leu Gly Ala Gly Phe Thr Gly Glu Ser Thr Phe Thr
 65 70 75 80
 Ser Lys Val Ser Ile Gln Phe Asn Leu Thr Ser Gly Met Thr Gly Leu
 85 90 95
 Arg Ala Ser Gly Asn Pro Ser Ile Ala Gly Phe Ser Gly Ile Ser Leu
 100 105 110
 Leu Ser Gly Leu Val Ala Glu Phe Cys Leu Glu Arg Pro Gly Ser Leu
 115 120 125
 Gly Leu Cys Ala Ile Tyr Ala Ala Trp Val Gly Gly Phe Ser Met Ser
 130 135 140
 His Arg Ser Met His Asp Phe Thr Arg
 145 150

<210> 201
 <211> 527
 <212> DNA
 <213> Homo sapiens

<400> 201
 gatgtggcta ttatccctgt ttcccagggtg agaaacaggg tcagtgatag agctggggtg
 60
 tgtgctgcga ggetcaccag ccagtccctt cctcaccaag gatgatgttc tccgtgggtga
 120
 gctggctcctt ggtctccttg aactcgtggc gcacctgggc cagctgcgcc tcgaaggcat
 180
 ccttctccat ctctttggtc agctgcaagt tctggagctg ctctgttgagg tctgtgatct
 240
 catccacctg ctggttgagc gtgcgcttga ggaaggccac aatctctctt ttgttattgg
 300
 ccagctgctc aaactccttg cggaaacatct tctcctgcac agccagctca tcccacttcc
 360
 gctgtgtaccg ggctagccgg tctctcagggt ctccgagctg gatgtggtag aactccttca
 420
 tctccttgcc cagaggcggc tccacggcca ccaccggctc cttcttgccc cttttcttct
 480
 tgacttcaag ctcttgctt gccttgctca cactcttttt gggaggc
 527

<210> 202

<211> 70
 <212> PRT
 <213> Homo sapiens

<400> 202
 Gly Arg Pro Gln Ser Pro Ser Cys Tyr Trp Pro Ala Ala Gln Thr Pro
 1 5 10 15
 Gly Gly Thr Ser Ser Pro Ala Gln Pro Ala His Pro Thr Ser Ala Gly
 20 25 30
 Thr Gly Leu Ala Gly Pro Pro Gly Leu Gly Ser Gly Cys Gly Arg Thr
 35 40 45
 Pro Ser Ser Pro Trp Pro Glu Ala Ala Pro Arg Pro Pro Ala Pro
 50 55 60
 Ser Cys Pro Leu Ser Ser
 65 70

<210> 203
 <211> 304
 <212> DNA
 <213> Homo sapiens

<400> 203
 ngtgcaccgg tggatcatgga caacgcgcgcc tacgtggtct acacctcggg atccaccggc
 60
 cgacccaagg gagttgtcgt caccacaccc gaactcgaca gcttcgcact cgaccagcag
 120
 cgtcgattcc acgcagatca ccaactctega accctgcact tcgccacccc cagcttcgac
 180
 ggagccgtct tcgagtacct gcaggcattc ggtgtcggag ccaccatggt gatcgccccg
 240
 accgacatct acggcgggcgc cgaactggca agtctcatcc gccgcgaaca cgtcactcac
 300
 gcgt
 304

<210> 204
 <211> 101
 <212> PRT
 <213> Homo sapiens

<400> 204
 Xaa Ala Pro Val Val Met Asp Asn Ala Ala Tyr Val Val Tyr Thr Ser
 1 5 10 15
 Gly Ser Thr Gly Arg Pro Lys Gly Val Val Val Thr His Thr Gly Leu
 20 25 30
 Asp Ser Phe Ala Leu Asp Gln Gln Arg Arg Phe His Ala Asp His His
 35 40 45
 Ser Arg Thr Leu His Phe Ala Thr Pro Ser Phe Asp Gly Ala Val Phe
 50 55 60
 Glu Tyr Leu Gln Ala Phe Gly Val Gly Ala Thr Met Val Ile Val Pro
 65 70 75 80
 Thr Asp Ile Tyr Gly Gly Ala Glu Leu Ala Ser Leu Ile Arg Arg Glu
 85 90 95
 His Val Thr His Ala

100

<210> 205

<211> 356

<212> DNA

<213> Homo sapiens

<400> 205

nngaattcag caatgataac tggctcaatt gaaggtgaaga caacaattga gggaattaat
 60
 gcacaattaa atacagtgtt aactttatatt tcaccacaat caaaagataa agatttaatt
 120
 atgccagatc aacaagaaga aatagatatt ctgattgcaa cgcactgtat ttcagaagga
 180
 cagaacttac aagattgtga ttacttaata aactatgaca ttcattggaa tccagttcgt
 240
 atcattcaaa gatttggacg gattgatcga attggttcga agaataaatg tgtacaatta
 300
 gttaactttt ggccagatat tacattagat gaatatattg atctaaaggg acgcgt
 356

<210> 206

<211> 118

<212> PRT

<213> Homo sapiens

<400> 206

Xaa Asn Ser Ala Met Ile Thr Gly Ser Ile Glu Gly Lys Thr Thr Ile
 1 5 10 15
 Glu Gly Ile Asn Ala Gln Leu Asn Thr Val Leu Thr Leu Phe Ser Pro
 20 25 30
 Gln Ser Lys Asp Lys Asp Leu Ile Met Pro Asp Gln Gln Glu Glu Ile
 35 40 45
 Asp Ile Leu Ile Ala Thr Asp Cys Ile Ser Glu Gly Gln Asn Leu Gln
 50 55 60
 Asp Cys Asp Tyr Leu Ile Asn Tyr Asp Ile His Trp Asn Pro Val Arg
 65 70 75 80
 Ile Ile Gln Arg Phe Gly Arg Ile Asp Arg Ile Gly Ser Lys Asn Lys
 85 90 95
 Cys Val Gln Leu Val Asn Phe Trp Pro Asp Ile Thr Leu Asp Glu Tyr
 100 105 110
 Ile Asp Leu Lys Gly Arg
 115

<210> 207

<211> 324

<212> DNA

<213> Homo sapiens

<400> 207

acgcgtgcac tgtgtgtatg catggtaacg tacacgtgtg cactgtgtgt ggtgtgcacg
 60
 catggtgtgt gcacgtgtng cactgtgtgt ggatgcatgg taatgtgcac gtgtgcactg
 120

tgtgtggtgt gtatgcatgg tgtgtgcacg tgtgcactgt gtgtgtgtgt atgcatgtgt
 180
 gtgcacatgt gcaactgtgtg gtgtgtatgc atggtgtgtg cacgtgtgca ctgtgtatgc
 240
 atngtgtgtg gcatgtgtgc actgtgtatg catagtgtgc acgtgtgcac tgtgtggtgt
 300
 gtatgcatgg taatgtgcac gtgt
 324

<210> 208

<211> 108

<212> PRT

<213> Homo sapiens

<400> 208

Thr Arg Ala Leu Cys Val Cys Met Val Thr Tyr Thr Cys Ala Leu Cys
 1 5 10 15
 Val Val Cys Met His Gly Val Cys Thr Cys Xaa Thr Val Cys Gly Cys
 20 25 30
 Met Val Met Cys Thr Cys Ala Leu Cys Val Val Cys Met His Gly Val
 35 40 45
 Cys Thr Cys Ala Leu Cys Val Cys Val Cys Met Cys Val His Met Cys
 50 55 60
 Thr Val Trp Cys Val Cys Met Val Cys Ala Arg Val His Cys Val Cys
 65 70 75 80
 Met Xaa Val Cys Met Cys Ala Leu Cys Met His Ser Val His Val Cys
 85 90 95
 Thr Val Trp Cys Val Cys Met Val Met Cys Thr Cys
 100 105

<210> 209

<211> 168

<212> DNA

<213> Homo sapiens

<400> 209

nnctccagag gttatgaggt tggaaagcccg gtttttttca ggtgcagaaa aggctaccat
 60
 attcaagggtt ccacgactcg cacctgcctt gccaatTTaa catggagtgg gatacagacc
 120
 gaatgtatac ctcatgcctg cagacagcca gaaaccccg cacaacgcg
 168

<210> 210

<211> 56

<212> PRT

<213> Homo sapiens

<400> 210

Xaa Ser Arg Gly Tyr Glu Val Gly Ser Pro Val Phe Phe Arg Cys Arg
 1 5 10 15
 Lys Gly Tyr His Ile Gln Gly Ser Thr Thr Arg Thr Cys Leu Ala Asn
 20 25 30
 Leu Thr Trp Ser Gly Ile Gln Thr Glu Cys Ile Pro His Ala Cys Arg

35 40 45
 Gln Pro Glu Thr Pro Ala His Ala
 50 55

<210> 211
 <211> 354
 <212> DNA
 <213> Homo sapiens

<400> 211
 tacatgggct ttgacacagt ggtggctgaa gctgcactaa ggggtgttgg aggcaatgtc
 60
 cagctggcag ctacagacct tgcacacccat ggaggaagcc tcccaccga cctgcagttc
 120
 tcaggagagg actcctcccc cacaccgtcc acatccccc ctgactctgc agggacctct
 180
 agtgccctga cagatgaaga catggagacg gaggctgtca acgaaatcct gggagacatt
 240
 ccggagcagc agggaggacta cctggactcc acgctggagg atgaagaagt cattattgtc
 300
 gaatacttgt cctgcgttga aagtataagt tctgccngca aagaacaact gatc
 354

<210> 212
 <211> 118
 <212> PRT
 <213> Homo sapiens

<400> 212
 Tyr Met Gly Phe Asp Thr Val Val Ala Glu Ala Ala Leu Arg Val Phe
 1 5 10 15
 Gly Gly Asn Val Gln Leu Ala Ala Gln Thr Leu Ala His His Gly Gly
 20 25 30
 Ser Leu Pro Pro Asp Leu Gln Phe Ser Gly Glu Asp Ser Ser Pro Thr
 35 40 45
 Pro Ser Thr Ser Pro Ser Asp Ser Ala Gly Thr Ser Ser Ala Ser Thr
 50 55 60
 Asp Glu Asp Met Glu Thr Glu Ala Val Asn Glu Ile Leu Glu Asp Ile
 65 70 75 80
 Pro Glu His Glu Glu Asp Tyr Leu Asp Ser Thr Leu Glu Asp Glu Glu
 85 90 95
 Val Ile Ile Ala Glu Tyr Leu Ser Cys Val Glu Ser Ile Ser Ser Ala
 100 105 110
 Xaa Lys Glu Gln Leu Ile
 115

<210> 213
 <211> 669
 <212> DNA
 <213> Homo sapiens

<400> 213
 attgcccaat ctacagagtgt ccaggaaagc ctggagagcc tttgtcagtc tattggggaa
 60

gttgaacaaa acctggaagg gaaacagggtg tcatactctt catcaggagt catccaggaa
 120
 gccttagcca caaatatgaa attgaagcag gacattgctc ggcaaaagag cagcttggag
 180
 gccaccctgt agatgggtgac ccgattcatg gagacagcag acagtactac agcagcagtg
 240
 ctgcagggca aactggcaga ggtgagccag cggttcgaac agctctgtct acagcagcaa
 300
 gaaaaggaga gctccctaaa gaagcttcta cccaggcag agatgtttga acacctctct
 360
 ggtaagctgc agcagttcat ggaacaacaa agtcggatgc tggcctcttg aaatcagcca
 420
 gatcaagata ttacacattt ctccaacag atccaggagc tcaatttga aatggaagac
 480
 caacaggaga acctagatac tcttgagcac ctggctcactg aactgagctc ttgtggcttt
 540
 gcgtggagt tgtgccagca tcaggacagg gtacagaatc taagaaaga cttcacagag
 600
 ctacagaaga cagttaaaga gagagagaaa gatgcattcat ctgccagga acagttggat
 660
 gaattccgg
 669

<210> 214

<211> 223

<212> PRT

<213> Homo sapiens

<400> 214

Ile Ala Gln Ser Gln Ser Val Gln Glu Ser Leu Glu Ser Leu Leu Gln
 1 5 10 15
 Ser Ile Gly Glu Val Glu Gln Asn Leu Glu Gly Lys Gln Val Ser Ser
 20 25 30
 Leu Ser Ser Gly Val Ile Gln Glu Ala Leu Ala Thr Asn Met Lys Leu
 35 40 45
 Lys Gln Asp Ile Ala Arg Gln Lys Ser Ser Leu Glu Ala Thr Arg Glu
 50 55 60
 Met Val Thr Arg Phe Met Glu Thr Ala Asp Ser Thr Thr Ala Ala Val
 65 70 75 80
 Leu Gln Gly Lys Leu Ala Glu Val Ser Gln Arg Phe Glu Gln Leu Cys
 85 90 95
 Leu Gln Gln Gln Glu Lys Glu Ser Ser Leu Lys Lys Leu Leu Pro Gln
 100 105 110
 Ala Glu Met Phe Glu His Leu Ser Gly Lys Leu Gln Gln Phe Met Glu
 115 120 125
 Asn Lys Ser Arg Met Leu Ala Ser Gly Asn Gln Pro Asp Gln Asp Ile
 130 135 140
 Thr His Phe Phe Gln Gln Ile Gln Glu Leu Asn Leu Glu Met Glu Asp
 145 150 155 160
 Gln Gln Glu Asn Leu Asp Thr Leu Glu His Leu Val Thr Glu Leu Ser
 165 170 175
 Ser Cys Gly Phe Ala Leu Asp Leu Cys Gln His Gln Asp Arg Val Gln
 180 185 190
 Asn Leu Arg Lys Asp Phe Thr Glu Leu Gln Lys Thr Val Lys Glu Arg


```

65          70          75          80
Gln Phe Asn Leu Phe Gly Ser Arg Thr Ile Tyr Asp Asn Val Ala Tyr
      85          90          95
Pro Leu Lys Leu Ala His Trp Lys Lys Ala Asp Glu Lys Lys Arg Val
      100          105          110
Thr Glu Leu Leu Ser Phe Val Gly Leu Thr Ser Lys Ala Trp Asp His
      115          120          125
Pro Asp Gln Leu Ser Gly Gly Gln Lys Gln Arg Val Gly Ile Ala Arg
      130          135          140
Ala Leu Ala Thr Lys Pro Ser Ile Leu Leu Ala Asp Glu Ser Thr Ser
      145          150          155          160
Ala Leu Asp Pro Glu Thr Thr Ala Asp Val Leu Ser Leu Leu Lys Arg
      165          170          175
Val Asn Ala Glu Leu Gly Val Thr Val Val Val Ile Thr His Glu Met
      180          185          190
Glu Val Val Arg Ser Ile Ala Gln Gln Val Ser Val Leu Ala Ala Gly
      195          200          205
His Leu Val Glu Ser Gly Ser Ala Arg Gln Val Phe Ala His Pro Gln
      210          215          220
Ser Glu Thr Thr Gln Arg Phe Leu Ala Thr Ile Ile Gly Gln His Pro
      225          230          235          240
Ser Gly Glu Glu Gln Ala Arg Leu Gln Ser Glu Asn Pro Asp Ala Arg
      245          250          255
Leu Val Asp Val Ser Ser Val Ala Ser His Ser Phe Gly Asp Ala
      260          265          270

```

<210> 217

<211> 500

<212> DNA

<213> Homo sapiens

<400> 217

```

nnacgcgctgc cgatgaaaga ggcgctgaaa ggtgccatcc agattccaac agtgactttt
60
agctctgaga agtccaatac tacagccctg gctgagttcg gaaaatacat tcataaagtc
120
tttctacagc tggtcagcac cagctttatc cagcatgaag tcgtggaaga gtatagccac
180
ctgttctacta tccaaggctc ggaccccagc ttgcagccct acctgtgat ggctcacttt
240
gatgtgtgtc ctgcccctga agaaggctgg gaggtgcccc cattctctgg gttggagcgt
300
gatggcgctc tctatggttg gggcacactg gacgacaaga actctgtgat ggcattactg
360
caggcccttg agctcctgct gatcaggaag tacatcccc gaagatcttt cttcattttc
420
ctggggccatg atgaggagtc atcaggggaca ggggctcaga ggatctcagc cctgctacag
480
tcaaggggag tccagctagc
500

```

<210> 218

<211> 166

<212> PRT

<213> Homo sapiens

<400> 218

Xaa Arg Val Ala Met Lys Glu Ala Leu Lys Gly Ala Ile Gln Ile Pro
 1 5 10 15
 Thr Val Thr Phe Ser Ser Glu Lys Ser Asn Thr Thr Ala Leu Ala Glu
 20 25 30
 Phe Gly Lys Tyr Ile His Lys Val Phe Pro Thr Val Val Ser Thr Ser
 35 40 45
 Phe Ile Gln His Glu Val Val Glu Glu Tyr Ser His Leu Phe Thr Ile
 50 55 60
 Gln Gly Ser Asp Pro Ser Leu Gln Pro Tyr Leu Leu Met Ala His Phe
 65 70 75 80
 Asp Val Val Pro Ala Pro Glu Glu Gly Trp Glu Val Pro Pro Phe Ser
 85 90 95
 Gly Leu Glu Arg Asp Gly Val Ile Tyr Gly Trp Gly Thr Leu Asp Asp
 100 105 110
 Lys Asn Ser Val Met Ala Leu Leu Gln Ala Leu Glu Leu Leu Ile
 115 120 125
 Arg Lys Tyr Ile Pro Arg Arg Ser Phe Phe Ile Ser Leu Gly His Asp
 130 135 140
 Glu Glu Ser Ser Gly Thr Gly Ala Gln Arg Ile Ser Ala Leu Gln
 145 150 155 160
 Ser Arg Gly Val Gln Leu
 165

<210> 219

<211> 361

<212> DNA

<213> Homo sapiens

<400> 219

acgcgttgaa acgggtatat tggggatgac gccgctgtgc aatatgcgca aggccataca
 60
 caaggtccgc acgctcccat gtccctcgtt ttgcacagtt cttttgcgcc gcattatggc
 120
 gaagccgtcg agattgcgcc tgatatcaag cgcacacgg tcaacaaccc cagccccc
 180
 acttttttgc gcaccaacag ttatctgate ggcgcgata cgctggcatt gatgatccc
 240
 ggtccgcttg acgagcccca tcacgcggcg ctgctgcgtg ccattgcggc cggccgggc
 300
 agccatatct ttgtcagcca cacacacogg gaccactcgc cagtgcgcac ggttttgaaa
 360
 g
 361

<210> 220

<211> 102

<212> PRT

<213> Homo sapiens

<400> 220

Met Ala Asp Arg Pro Ala Gly Asn Gly Thr Gln Gln Arg Arg Val Met

```

      1           5           10           15
Gly Leu Val Lys Arg Thr Gly Ile Asp Gln Cys Gln Arg Ile Ala Ala
      20           25           30
Asp Gln Ile Thr Val Gly Ala Glu Lys Ser Glu Gly Ala Gly Val Val
      35           40           45
Asp Arg Asp Ala Leu Asp Ile Arg Arg Asn Leu Asp Gly Phe Ala Ile
      50           55           60
Met Arg Arg Lys Arg Thr Val Glu Asn Glu Gly His Gly Ser Val Arg
      65           70           75           80
Thr Leu Cys Met Ala Leu Arg Ile Leu His Ser Gly Val Ile Pro Asn
      85           90           95
Ile Pro Val Ser Thr Arg
      100

```

<210> 221

<211> 401

<212> DNA

<213> Homo sapiens

<400> 221

```

agatctctgt gtcgtcggct gcaaagagga tgagcccaga tgcataatcag gggctccctc
60
ccacatccca cctgctcggg cagcccaagg cagcccccaca ctgctgcagc acacctcgct
120
gcagctctgg ttctctctca gaaatatccc tgccaccctg ctaagccttg gccaacactg
180
caccctgtcc caatgcggct ccagtgaacca cccccccagg gcataccctc ctacagagca
240
ttcccaaaaa aggctagagt agacaccagc ctgctccgta gggggcctcc accccattct
300
ccaaggcctc caccagggga cgctgggtga accgacatcc aggcttgccc caccctccctg
360
ctcagagtcc atgttctgtg acaaggggtg caactgggat t
401

```

<210> 222

<211> 124

<212> PRT

<213> Homo sapiens

<400> 222

```

Met Asp Ser Glu Gln Gly Gly Gly Pro Gly Leu Asp Ala Gly Ser Pro
      1           5           10           15
Gly Val Pro Gly Trp Arg Pro Trp Arg Met Gly Trp Arg Pro Pro Thr
      20           25           30
Glu Gln Ala Gly Val Tyr Ser Ser Leu Phe Trp Glu Cys Ser Val Gly
      35           40           45
Gly Tyr Ala Leu Gly Val Trp Ser Leu Glu Pro His Trp Asp Arg Val
      50           55           60
Gln Cys Trp Pro Arg Leu Ser Arg Val Ala Gly Ile Phe Leu Arg Arg
      65           70           75           80
Asn Gln Ser Cys Ser Glu Val Cys Cys Ser Ser Val Gly Leu Pro Trp
      85           90           95
Ala Ala Arg Ala Gly Gly Met Trp Glu Gly Ala Pro Asp Met His Leu

```


100 105 110
 Gly Ser Ser Ser Leu Gln Pro Thr Thr Gln Arg Ser
 115 120

<210> 223
 <211> 331
 <212> DNA
 <213> Homo sapiens

<400> 223
 tcatgaaatc tgtgggcagt gaccaggag ggtatgggca ggcccaacca ggttggtgtg
 60
 cccttgaagc cccacagacc tgccagggca gcagggcagt tgggagccgg agaacctgag
 120
 aaccaagcca ggtgcatgc aggaggctgg cacgtgaacg ctgcaggtgt tgccggcagc
 180
 cgtggtgcct ggcagatagt gtctgacccc cnaggacett cttgctgggc agcccagctc
 240
 aaaagctgtt cccgcttaag ccacccccac cgccttggcc acacctggca catgggtgaa
 300
 gcaaggggcat ttcccggggt ttctgtttcc c
 331

<210> 224
 <211> 103
 <212> PRT
 <213> Homo sapiens

<400> 224
 Met Pro Leu Leu His Pro Cys Ala Arg Cys Gly Gln Gly Gly Gly Gly
 1 5 10 15
 Gly Leu Ser Gly Asn Ser Phe Trp Thr Gly Leu Pro Ser Lys Lys Val
 20 25 30
 Leu Gly Gly Arg Thr Leu Ser Ala Arg His His Gly Cys Arg Gln His
 35 40 45
 Leu Gln Arg Ser Arg Ala Ser Leu Leu His Ala Ala Trp Leu Gly Ser
 50 55 60
 Gln Val Leu Arg Leu Pro Thr Ala Leu Leu Pro Trp Gln Val Cys Gly
 65 70 75 80
 Ala Ser Arg Ala His Gln Pro Gly Trp Ala Cys Pro Tyr Pro Pro Gly
 85 90 95
 Ser Leu Pro Thr Asp Phe Met
 100

<210> 225
 <211> 339
 <212> DNA
 <213> Homo sapiens

<400> 225
 tgatcacggg cgtgagccac cagcccagca tcccttgctt ttcattcgca cctccacetc
 60
 cagaatgacc etcattccct cctgcacaga cgggtgacagc agtaactcct acaaacacca
 120

ccagactgat cttcaagagc agaggaactc ccaatcacga ttccaccccc gccgggctct
 180
 caaatctctc aggggtgcct gctatggggg agggaggcac actttgtgtg gctctcaagg
 240
 cctcagccag ccgggtccaa accaaactcc agcctggcct caccatccca ccgccaaaac
 300
 tttgctcaca ctggcccttc ttcttggaac atgggcctn
 339

<210> 226

<211> 91

<212> PRT

<213> Homo sapiens

<400> 226

Met	Thr	Leu	Ile	Pro	Ser	Cys	Thr	Asp	Gly	Asp	Ser	Ser	Asn	Ser	Tyr
1				5					10				15		
Lys	His	His	Gln	Thr	Asp	Leu	Gln	Glu	Gln	Arg	Asn	Ser	Gln	Ser	Arg
			20					25				30			
Phe	His	Pro	Arg	Arg	Ala	Leu	Lys	Ser	Ser	Arg	Ala	Ala	Cys	Tyr	Gly
			35				40				45				
Gly	Gly	Arg	His	Thr	Leu	Leu	Gly	Ser	Gln	Gly	Leu	Ser	Gln	Pro	Gly
	50				55				60						
Pro	Asn	Gln	Leu	Pro	Ala	Trp	Pro	His	His	Pro	Thr	Ala	Lys	Pro	Leu
65			70					75					80		
Leu	Thr	Leu	Ala	Pro	Leu	Pro	Gly	Thr	Trp	Ala					
			85					90							

<210> 227

<211> 353

<212> DNA

<213> Homo sapiens

<400> 227

gtcgaccctc tcgattgtgg cgaactccat ggctgctgcg ggcctgcgta ggctctcgag
 60
 tagctcgacg tcgggttcgc gagggctcgc agcgtggcca tgctgtctct tggatggttc
 120
 gggcaactcc tcgggggatt cgagcagttc ttggcgccac tgctctggcg tcatccggga
 180
 ggccaggccg acaagtgtcg cctcctgcc aacgctgagc gacgctgcc tgttgagtag
 240
 ggcgtcttca ctggtcaggg cgagcgcgg atcgaccagg ttggcgccca ggccgagaga
 300
 cagcatgtct gctcagtcgc ggtgatgact ggagtggcgg tctcctgcac ggg
 353

<210> 228

<211> 102

<212> PRT

<213> Homo sapiens

<400> 228

Met Leu Ser Leu Gly Leu Asp Ala Asn Leu Val Asp Thr Ala Leu Ala